

How affordable is public health?

The only certainty about Britain's next general election campaign (postponed until next year) is that the argument will hinge on how the public health service is run and paid for. But the real issues are likely to be hidden in noise.

BRITAIN'S National Health Service (NHS), often known in the United States as "socialized medicine", has been a powerful influence for civility since its creation in 1947. Its founder, the late Aneurin Bevan, may have miscalculated in reckoning that a comprehensive health service ready to treat all comers at no cost would so improve public health that it would be a declining financial burden, but the benefits of the arrangement have been huge. Not the least of them is that children, the chronically sick, the indigent and the elderly have been well cared for, whatever their circumstances. Unseemly pauperization of the terminally ill by medical bills (or by the cost of insuring against them) has been outlawed.

How can such a laudable institution have become politically contentious? One explanation is Bevan's miscalculation; the global cost has not fallen, but has instead risen. This year, the cost to the British exchequer will exceed £33,000 million, one seventh of total public spending. Counting the cost of private medicine and the not insubstantial charges levied by the NHS on prescriptions and other services, British spending on health is approaching 10 per cent of Gross Domestic Product (GDP). This may be less than the proportion of GDP spent on medical services in West Germany, and is certainly much less when measured as spending per head of population, but the high cost and the certainty that it will continue to increase is calculated to alarm governments.

The more immediate explanation is the reorganization of the NHS embarked on four years ago, after a formal but private government review of the finances of the NHS. Ministers shrank from shifting part of the burden from taxpayers to working adults (and their employers) by means of medical insurance, settling instead for what is called "an internal market" in which distinguishable units within the NHS buy and sell services to each other. In this spirit, a number of physicians' practices are now being given annual budgets with which to buy medical services (from hospitals, for example) on behalf of their patients, while several hospitals have been given the status of self-managing trusts which, although still publicly owned, must balance their books financially by contracting with health authorities (the geographical units of the NHS) for the provision of services.

The bone of contention to emerge during the past two weeks centres on the opposition Labour Party's charge, at its annual conference, that the government has it in mind

to transfer substantial parts of the NHS to the private sector, and that the reorganization is a prelude thereto. Last week, both the Secretary of State for Health, Mr William Waldegrave, and the Prime Minister, Mr John Major, found it prudent to give formal public promises that they intend nothing of the kind. They are now in the position of those who attempt to answer questions such as "When did you stop beating your wife?": the presumption of guilt denied will not easily be shaken off. Party politics apart, the emergence of this row is potentially a misfortune for the NHS and for its patients. The obvious danger is that the reorganization still under way will be shot through with further uncertainty. It is hazardous enough already. An internal market, even if it works smoothly, will not of itself influence the total demand for medical services. Rather, the calculation is that financially free-standing units such as hospitals will have an incentive to function as efficiently as they can. In reality, the new status of the trust hospitals is much like that of the many public research laboratories (in defence, for example) converted into 'agencies' in the past few months, except that the hospitals can in principle play off one health authority against another when seeking business, whereas research laboratories usually have only one customer. Even so, the government must be wishing it had had more experience of this novel form of public administration before embarking on its reorganization of health.

There are many practical difficulties. One is that hospitals have no fund of experience on which to base the prices charged for services. Another is that one route to low-cost efficiency — specialization — will make them seem to patients to be neglecting their pastoral and community responsibilities. Yet the financial troubles already encountered by some London hospitals could be taken as a sign that the internal market is working: London is better supplied with hospitals than other parts of Britain, so that prospective financial deficits may be a sign that not all of them are viable. If that were true, the casualty would be medical teaching, traditionally concentrated in London and now inadequately provided for.

But the overarching problem is that no amount of saving brought about by greater internal efficiency can permanently halt the rising cost of up-to-date medical treatment. It is not merely that people who live longer have more opportunity to fall sick, but that medicine's ambitions for the treatment of patients have enormously



expanded. Finance ministers everywhere are no doubt glooming at the steady improvement of therapies that prolong the lives of cancer or AIDS patients, for example. What they will have to recognize is that there is nothing sacrosanct about present proportions of GDP spent on medicine and that, a few decades from now, they will seem small. Then the question of how to meet the cost will re-emerge. □

Aquarium Fights Back

The New England Aquarium, accused by animal rights groups of violating regulations, is countersuing for defamation.

FOR years, the New England Aquarium in Boston, known for coming to the rescue of beached whales, dolphins and seals, has been fighting legal assaults from animal rights groups who charge aquarium scientists with illegally taking animals from the sea. The most noted dispute centers on a dolphin named Kama that was transferred four years ago from the aquarium to a US Navy base for sonar research.

The animal rights groups say the transfer was illegal. Under the US Marine Mammal Protection Act, an aquarium must get a government permit to keep (or transfer) any animal taken from the wild. Because the aquarium had no permit, it had no right to send Kama to the Navy where, the groups claimed, Kama would be turned into a "lethal dolphin torpedo". The transfer was especially egregious, they contend in a flier designed as a fund-raiser for their groups, poor Kama was "violently removed from his family group and put into captivity at the New England Aquarium".

Not so, say aquarium scientists. To begin with, Kama was born not in some romantic natural spot in the ocean near Bermuda; he was born at Sea World, an aquarium in San Diego. (The rights groups dismiss this fact as a "minor mistake".) Further, aquarium scientists insist that the Navy promised not to use Kama for other than sonar studies and report that an independent oversight panel confirms this.

What finally pushed the aquarium over the edge? What drove a conservative institution to risk a countersuit in court, with all the legal costs entailed? In June three animal rights filed a suit against the aquarium alleging that it has failed to obtain necessary permits *before* making any attempts to rescue beached mammals. Obtaining such permits could take up to six months, aquarium officials state, and the delay would "guarantee certain death to these animals".

Arguing that the animal groups real agenda is to put all aquariums out of business, the New England Aquarium now threatens to go to court to recover \$5 million in damages for what it says are false and defamatory statements against it. The aquarium alleges that the misleading stories about Kama have caused some of its patrons to drop their contributions, while the coffers of the rights groups have increased.

It is high time that institutions under attack from extremist animal rights groups fight back. A dozen or so aquariums in the United States are expected to file amicus or "friend of the court" briefs on behalf of the New England Aquarium, which is a good sign of solidarity. It is also proper that the aquariums have decided to fight harassment with harassment. Should the New England Aquarium win its case (and \$5 million), animal rights extremists would get a deserved dose of their own medicine. As the aquarium points out, it costs millions to fight lawsuits in the courts and scientists are tired of seeing resources go to lawyers that could otherwise go to research.

Even more important, the aquariums should use this occasion to engage in a public debate on the real issues. No sensible person believes that the animal rights groups are really concerned about permits. They want to change laws that allow aquariums to keep and display sentient mammals under any circumstances. Their argument that human beings have no business keeping intelligent sea mammals in captivity is not without merit and deserves to be debated.

But, as the New England Aquarium points out, a legitimate argument can be made in favor of displaying sea mammals on educational grounds. The aquarium argues, for instance, that the one million people who visit every year (150,000 of them are school children) benefit greatly from seeing mammals in the aquarium's pools and that, by extension, mammals in the wild benefit too. Aquarium scientists want public support (both financial and moral) for studies of marine habitats and the preservation of populations of whales, penguins and other mammals. They are particularly interested at present in preserving the endangered North Atlantic right whale which may be dying off as a result of inbreeding.

How, scientists ask, can the general public understand the importance of the research at the New England Aquarium (and others) if people never have an opportunity to see a whale or a dolphin or a seal up close? Speaking of inner city children, the aquarium director put it vividly in an interview when he said, "The only animals they see are rats, pigeons, cats and dogs. If that's all you see, why should you be interested in conservation?" If they are taught that scientists steal cute dolphins from their homes and turn them over to the military, children surely will not think well of science.

It is no exaggeration to say that, in this and other disputes between the scientific community and animal rights groups, the issue goes to the very heart of the value of research on animals in all its many forms. Particularly in the aquarium's case, this is not a dispute over nuances. It is about the ethics of using animals for human benefit. It is imperative that animals be treated humanely, but is also reasonable that researchers themselves be treated humanely—that is, not subject to years of petty harassment that is merely a ploy for avoiding debate on the real issue. The aquarium is right to go to court. □

Europe approves first transgenic animal patent

- Decision could spur EC patenting
- Activists mobilise to challenge ruling

Brussels

THREE years after the US Patent Office granted the world's first patent for a genetically engineered animal — Harvard University's transgenic 'oncomouse' — the European Patent Office (EPO) has decided to go the same way. Although opponents of animal patenting promise to challenge the EPO ruling, the move will give added impetus to a European Commission directive, now being considered by the European Parliament, that would legalize animal patenting in the 12 European Communities (EC) states (See *Nature* 353, 8; 5 September 1991).

The oncomouse was developed in 1984 by Philip Leder and Timothy Stewart, then both at Harvard University. It carries several genetic sequences containing mouse oncogenes and promoter regions, and so is extremely susceptible to tumours, making it a useful model for studying cancer. But Harvard's patent application was worded sufficiently broadly to cover all transgenic mammals containing activated oncogenes. The new European patent applies to 11 of EPO's 14 member countries, including the influential 'big four' EC members — Britain, France, Germany and Italy — plus four other EC nations.

While the US patent office moved fairly quickly to grant a patent, Harvard's European application has been the subject of bitter controversy, and has languished in EPO's Munich office since 1985. In 1989, EPO patent examiners rejected the application, arguing that the European Patent Convention prohibits the patenting of animals (see *Nature* 340, 85; 13 July 1989). But after a successful appeal by Harvard, the examiners were last year told to reconsider their decision, and to weigh the benefits to cancer research of granting the patent against the likely suffering of the tumour-susceptible animals themselves.

Opponents of the decision now have nine months in which to issue a formal challenge. "The case is very likely to continue," says EPO spokesman Rainer Osterwalder. And, indeed, a coalition of European animal welfare groups, headed by the British organization Compassion in World Farming, is already planning to fight the decision.

Compassion in World Farming is concerned that the oncomouse ruling will clear the way for the patenting of transgenic farm animals. Among a number of transgenic animal patent applications be-

ing reviewed by EPO is a request from Merck and Co. for a patent on a chicken carrying the gene for bovine growth hormone. Compassion in World Farming's legal director Peter Stevenson argues that such a rapidly growing chicken is likely to suffer excessively from leg and foot deformities.

Stevenson's challenge will be made under an article of the European Patent Convention that can deny a patent if the invention is deemed to be "contrary to public order or morality". He believes that concern over the welfare of transgenic animals justifies a challenge on moral grounds. But Norman Davies, head of patents with the British Technology Group, says he knows of no case in which the 'public order' clause has been used successfully to challenge a patent.

Despite the probable challenge, the EPO decision to grant the oncomouse patent will boost the prospects of a European Commission directive on biotechnology patents that would make legal the patenting of transgenic animals throughout the EC. First proposed in 1988, the directive has struggled to find time on the European Parliament's agenda, and was recently rejected by the Parliament's Agriculture Committee.

But Commission biotechnology officials are now optimistic that the Parliament's Legal Affairs Committee, which has the lead role in advising the full Parliament on the issue, will recommend that the directive be accepted. That committee is due to vote at the end of this month, and is expected to allow the patenting of transgenics, although it may ask for some restrictions to ensure animal welfare.

Assuming the directive passes, the full Parliament may vote on the directive as early as next month, after which the language would be sent to ministers of the EC member states to be formally adopted.

Commission biotechnology officials hope the latest developments in the Parliament, together with the new EPO decision, will provide the right climate to press forward quickly with the new directive. The directive contains a range of provisions to harmonize the patenting of biotechnology products across the EC and, with the 'single European market' just over one year away, Commission officials would like to see the directive enforced as soon as possible.

Peter Aldhous

Green light for ddl

Washington

AFTER an expedited review of just six months, the US Food and Drug Administration (FDA) gave the New York-based drug company Bristol-Myers Squibb approval to market its antiretroviral drug Videx (also known as dideoxyinosine or ddl) for the treatment of AIDS. The drug, approved initially for narrow use, can be used to treat both adult and paediatric AIDS patients who are intolerant to AZT, the only other approved drug for the treatment for HIV infection.

Under an expanded access programme and through clinical trials, the company estimates that more than 30,000 patients have already received VIDEX therapy. After a review of preliminary phase I trial data, the drug was approved on the basis of ddl's effect on increasing the patient's CD4 helper cell count, a surrogate marker for determining clinical effectiveness.

Donald Abrams, an AIDS researcher at the San Francisco General Hospital and a member of FDA's antiviral drug products advisory committee which recommended approval of Videx, says that while ddl clearly has biological activity, he would be happier if HIV patient survival rates were as encouraging as their improvement in CD4 counts after treatment with the drug.

The company, which estimates that a year's supply of Videx will cost about \$1,745, has set up a reimbursement assistance programme to help patients identify sources of third-party reimbursement.

Diane Gershon

Challenge to AZT

New Delhi

INDIA is the latest country to challenge the Burroughs Wellcome monopoly on AIDS drug AZT. Three Indian pharmaceutical companies expect to sell the drug on the domestic market and for export before the year's end.

The Indian Institute of Chemical Technology (IICT), which developed the process for synthesising AZT, has licensed the process free of cost to Lupin, Cadila and Reddy Laboratories. According to Dr. A.V. Rama Rao, director of IICT, the Indian-made AZT will cost a quarter or less of the current Wellcome prices.

The fact that AZT has been at the centre of patent disputes in the United States and Canada does not affect India, which has not signed the Paris convention and does not recognize patent protection for pharmaceuticals. Under Indian patent law, only processes, and not products, may be patented. Several Indian companies are exploring the foreign market in Africa and other areas where Burroughs Wellcome has no patent. The World Health Organization has shown interest in the Indian product and may become the largest buyer, Rao says.

K.S. Jayaraman

Integrity review halted

Washington

LAST July, while congressional investigators were grilling the US National Institutes of Health for their failing scientific misconduct policies, another biomedical agency — the Food and Drug Administration (FDA) — was quietly starting a review of its own integrity procedures, hoping to catch the flaws before Congress did.

But now the FDA review has itself run into trouble. Last month, just nine weeks after the agency issued a \$710,558 contract to a Washington law firm to run the review, it was forced to cancel the contract when the congressional investigative committee of Representative John Dingell (Democrat, Michigan) launched an investigation into possible conflict of interest at the winning firm, Washington-based Shaw, Pittman, Potts & Trowbridge.

Now, on top of a continuing probe into data-integrity lapses that have allowed drug companies to submit spurious products and clinical trial results (typified by last year's generic drug scandal), the Dingell committee is investigating how the agency could have awarded the integrity-review contract to a law firm that was simultaneously representing major FDA clients and had bragged that the contract would help it to get those clients through the system.

According to contract documents, FDA had decided to bring in outside help because staff and resource shortages have forced the agency "to rely more heavily on entities submitting data to it to police themselves" — a system that, FDA admitted, "has not always worked as well as it should." The agency receives more than 150,000 submissions for new drugs, medical devices, food additives and the like each year, and most applications are accompanied by supporting data to demonstrate safety and efficacy. In some cases — as an FDA investigation of a faulty artificial heart valve last year proved — those data are sometimes intentionally incomplete, misleading, and even fabricated.

FDA hoped that going to a private law

firm for an "integrity review" might convince Dingell and other congressional watchdogs that it was serious about reforming itself. But picking Shaw, Pittman as that firm appears to have only made the matter worse.

In an article published on 2 September, the trade journal *Legal Times* revealed that, far from being conflict-free, Shaw, Pittman has major pharmaceutical and food companies (including Monsanto, Dow Chemical and RJR Nabisco) as its clients. J. Thomas Lenhart, one of the partners, told *Legal Times* that investigating FDA's data integrity procedures would "give us a chance to learn more about a substantial area of legal work....It may well open up new professional opportunities in the future."

The firm claimed that, because it lobbied for the companies, rather than directly handling their FDA applications, there was no conflict. But Dingell was unpersuaded, especially when he learned that the firm was also behind a new group — known as the Workplace Health and Safety Council — intended to oppose the strengthening of occupational health and safety laws. "It's a hugely embarrassing mess," says one Dingell staff member.

Within a week of the article, the Dingell committee had sent a letter to FDA, requesting all documentation on the contract awarding process and the integrity review. Committee staff then interviewed the FDA officials in charge of the review project. On 30 September, FDA gave in. It cancelled the contract and announced that it would finish the review internally.

Although Dingell staff have not planned a hearing on the case, they are looking into both the handling of the contract and the relationship between FDA and Shaw, Pittman. The case "exposed the vagaries of Washington deal-making," says a staff member. Whether that went as far as "an arrangement to give FDA a Good House-keeping seal of approval", the staffer adds, will be the focus of the continuing investigation.

Christopher Anderson

Social science gets its day

Washington

AFTER years of lobbying for their own directorate at the US National Science Foundation (NSF), social and behavioural scientists have finally been granted their wish. NSF last week announced that it is ending the second-class status of these disciplines in the current directorate for Biological, Behavioral, and Social Sciences (BBS) by splitting off a new \$70 million Social, Behavioral and Economic Sciences directorate. Completing the name

game, BBS will become simply the directorate for Biological Sciences. NSF director Walter Massey said he hoped that having a directorate of their own will mean "more clout" for both disciplines, but did not announce a major funding increase in either area. The Consortium of Social Science Association last week hailed the changes as a long overdue effort to remove social and behavioural sciences from the "long shadow" of biology at NSF.

Christopher Anderson

Healy back on the beat

Washington

HOPING to speed misconduct reforms at the US National Institutes of Health (NIH), assistant secretary for health James Mason last week returned NIH director Bernadine Healy to the driver's seat. Healy had withdrawn herself from decisions affecting the NIH Office of Scientific Integrity (OSI) after it emerged that OSI was investigating a researcher who had worked under her at her former institute, the Cleveland Clinic Foundation (see *Nature* 352, 461; 8 August 1991).

Explaining that the six to seven months remaining of the Cleveland Clinic investigation was too long to be without Healy's help, Mason transferred the case to the Office of Scientific Integrity Review (OSIR) in NIH's parent agency, the Department of Health and Human Services.

Just how OSIR will handle the case is not clear; the office, which is set up to review OSI investigations and assign penalties in a trial-like setting, has no investigators of its own, nor any experience in the art of hands-on investigating. Predictably, the investigations committee of Representative John Dingell (Democrat, Michigan) is planning a letter of protest.

C.A.

WIPP put on hold

Washington

THE US Department of Energy (DOE) has backed down from a decision to begin shipping nuclear waste to its Waste Isolation Pilot Plant (WIPP) in New Mexico after the state's attorney general filed a lawsuit to stop the shipment. DOE had announced two weeks ago (*Nature* 353, 487; 10 October 1991) that WIPP was ready to be put into operation and that the energy department would not wait any longer for Congress to approve legislation on the last step needed to open the waste site — the transfer of the land to DOE.

But last week, after the state of New Mexico and environmental activists filed lawsuits, DOE reversed itself. Halting the process until a US court hears the New Mexico case on 15 November, "will help ensure that legal issues are resolved prior to our first shipment from the Idaho National Engineering Laboratory," Secretary of Energy James Watkins said.

Watkins was more conciliatory last week than the week before, when he had threatened to cancel \$63 million in payments to the state if it filed suit; now he has assured the New Mexico governor of "the department's financial and environmental commitments to the State of New Mexico".

Robert Pool

Cracks in the RAC

Washington

LAST week, the US National Institutes of Health recombinant DNA advisory committee (RAC) approved what is billed as the first experiment in human gene therapy designed to immunize a cancer patient against his own tumour. But it rejected out of hand the recommendation of its own gene therapy subcommittee that it approve a protocol for the treatment of ovarian cancer with a genetically engineered "vaccine", a move that cast new doubts on the future of the subcommittee.

The RAC unanimously approved two protocols submitted by Steven Rosenberg of NIH's National Cancer Institute (NCI) that will be the first attempt to use gene therapy to find a cancer vaccine for melanoma. With no shortage of patients, Rosenberg began treatment on a 46-year-old man with metastatic melanoma within 24 hours of receiving RAC approval, followed immediately by approval from NIH director Bernadine Healy. Rosenberg first introduced the gene coding for tumour necrosis factor (TNF), a potent anti-tumour toxin, into a tumour cell line established from the patient. The gene-modified tumour cells were then injected back into the patient in the hope that they will be more immunogenic than unaltered tumour cells. In order to boost the hoped-for anti-tumour effect, lymphocytes will be withdrawn after 21 days from the patient, expanded or multiplied in the laboratory, and then be reinfused into the patient. Rosenberg's second protocol, almost identical in detail to the first, employs interleukin-2 instead of TNF. The RAC also approved the experimental therapy for patients with advanced colorectal and renal carcinomas.

But the ovarian cancer protocol, which was given the green light by the subcommittee, was shot down by the RAC in a 19 to 0 vote. In earlier *in vitro* experiments and in mouse studies, Scott Freeman of the University of Rochester Medical Center had demonstrated that transfer of the gene for herpes simplex virus-thymidine kinase into ovarian tumour cells renders the cells more sensitive to antiviral agents such as ganciclovir. Freeman had planned to apply this rationale to the treatment of patients with ovarian cancer who were in relapse and who had already received chemotherapy and radiation treatment.

William Kelley of the University of Pennsylvania Medical Centre, who sits on both the subcommittee and the RAC and who was a primary reviewer of the Freeman protocol, had strong reservations about the feasibility of the study but left before the vote was taken at the July subcommittee meeting. Kelley says he felt that a lot more work needed to be done before approval would be warranted and

admits to being "very surprised" that approval had been granted.

Disagreement about the protocol centres on whether Freeman has obtained sufficient data from animal studies.

In addition to the focus on specific gene therapy protocols, the meeting generated a rising tide of opinion among committee members that the two-tiered review process (where protocols must first pass the subcommittee, then the full RAC) is unwieldy, involves considerable duplication of effort and sends confusing signals to researchers. Phasing out the human gene therapy subcommittee in the long term now seems inevitable.

In the early 1980s, the RAC had responsibility for a number of research areas, including experiments that involved the deliberate release of genetically engineered organisms. But as concerns about recombinant organisms have diminished, the purview of the RAC has diminished and its role has become almost parallel to that of the gene therapy subcommittee. The subcommittee was set up in the summer of 1984 in anticipation of a flood of research protocols involving somatic-cell gene transfer and gene therapy. Today, the RAC evaluates little else.

As someone who speaks from personal experience, having successfully steered one of the first protocols for gene therapy through a tortuous review process, W. French Anderson of NIH's Heart Lung and Blood Institute is an outspoken critic of the two-tiered NIH system. "The point is", Anderson says, "there really is no need for two separate committees" because the RAC itself now has time to review protocols in detail without prior subcommittee review. Anderson, with Rosenberg and Michael Blaese of NCI were the first to conduct a trial of gene transfer when they used a genetically engineered marker gene to track anti-tumour lymphocytes in patients with melanoma.

Henry Miller, who heads up the Office of Biotechnology within the FDA, agrees with Anderson and favours consolidation of the two committees. "Given that the full RAC has very little to occupy it other than gene therapy, it seems self-evident that it should and can reconstitute itself to be primarily concerned with gene therapy", says Miller, whose agency also reviews human gene therapy protocols.

In the short term, the RAC is attempting to streamline the review process to make it more efficient. And, although change is not expected overnight, a subcommittee has been appointed to consider whether the gene therapy subcommittee and the full committee should be merged. The new subcommittee on mergers has until February to reach an opinion.

Diane Gershon

Approvals still slow, but pipeline full

Washington

ONE hundred and thirty-two drugs and vaccines are now in development at US biotechnology companies, a 63 per cent increase since 1988, according to a report released last week by the US Pharmaceutical Manufacturers Association (PMA) entitled "Biotechnology Medicines in Development". Of these, half are being tested for cancer, or cancer-related conditions.

But although the future clearly looks bright, biotechnology in the present is still more smoke than fire. Just three biopharmaceuticals have been approved by the US Food and Drug Administration (FDA) in the last 16 months, the date of the last PMA survey, and only 14 such biotechnology products have been approved over the last 9 years. Given this disturbing trend, the PMA has estimated that, at current staffing and funding levels, it would take the FDA 13 years to approve the 21 medicines already awaiting approval.

The US patent system is equally to blame in slowing biotechnology developments, PMA found. Although the patent office initiated a programme in 1988 to reduce the processing time of biotechnology patent applications, the average processing time of 25.6 months still lags far behind the 18-month average for all patent applications. Much of the reason for this is no doubt the 15 per cent annual growth rate in filings for biotechnology inventions — more than twice the average annual increase of 7 per cent in all patents.

The United States can, however, draw comfort from the fact that of the 169 genetic engineering health-care patents issued in 1990, 82 per cent were of US origin; Japan was second with 11 per cent.

D.G.

ANTARCTIC OZONE

Worst hole yet

Washington

STRATOSPHERIC ozone levels over Antarctica have reached the lowest levels on record, researchers monitoring a National Aeronautics and Space Administration (NASA) satellite reported last week. In its most recent pass over Antarctica, the Total Ozone Mapping Spectrometer (TOMS) aboard NASA's Nimbus 7 satellite recorded ozone levels of less than 120 Dobson units (a measure of atmospheric opacity to ultraviolet and other light), and tentative measurements as much as 10 units lower. That compares to measurements of 500 units in normal conditions. "The minimum ozone on October 6, 1991 [a 110 Dobson unit reading], is the lowest we have ever seen with the TOMS instrument in its 13-year record of data," said NASA researcher Arlin Krueger. C.A.

Maglev burns out in Japan

Tokyo

JAPAN'S ambitious plans to develop a high-speed magnetically levitated (maglev) train went up in smoke — literally — on 3 October, when a test vehicle burnt to the ground during a test run. The accident throws into question the future of maglev trains, a technology that Japan is hoping to develop into its commercial transport system of the future.

The 22-metre maglev train is capable of travelling at more than 400 km an hour when floating on its superconducting magnets, but the accident occurred when the vehicle was only running at low speed on rubber tyres. One of the tires went flat and burst into flames, and confinement of the fire by the concrete walls of the U-shaped trackway resulted in the train being completely gutted.

Although three researchers on the train escaped unhurt, the fire raises serious questions about the safety of the maglev system. From 1993, the Railway Technical Research Institute, which developed the maglev train, plans to start testing it on a new 43-km track that will run through long tunnels in the mountains of Yamanashi Prefecture west of Tokyo. (Until now, the train has been tested only on a 7-km overhead trackway in Japan's southern island of Kyushu where the accident occurred). If a similar fire broke out in a tunnel it could have disastrous consequences.

Fire is not the only problem the maglev train faces. Researchers are having difficulty shielding the inside of the vehicle from the powerful magnetic fields emitted from the superconducting magnets that drive and levitate the train. At present,

people with heart pacemakers cannot travel in the train for fear that the pacemaker might be stopped by the magnetic fields. And the medical effects of long-term exposure to strong magnetic fields are also uncertain.

Another potential danger is 'quenching' of the superconducting magnets, a process (albeit rare) whereby the magnets rapidly lose their superconductivity. The maglev train in such an event would drop onto its rubber tyres, and the present accident suggests that those cannot be relied on.

Despite these problems, Transport Minister Kanezo Muraoka, whose ministry is financing the project, says tests on the new track in Yamanashi Prefecture will go ahead on schedule. The project has the backing of one of Japan's most powerful politicians, Shin Kanemaru, the so-called 'godfather' of the ruling Liberal Democratic Party who is reputed to have the power to select Japan's prime ministers and whose constituency is in Yamanashi Prefecture. Although replacing the burnt-out vehicle will cost about ¥3,000 million (\$23 million), this cost is small compared with the potential financial and political fallout of cancellation.

Property speculators have already moved into the Yamanashi area and driven up the prices of land near the test track, and too much is now at stake politically for the project to be called off. But ambitious plans by proponents of the maglev train to extend the Yamanashi trackway into a commercial line linking Tokyo and Osaka must now be very much in question because of the numerous technical hurdles in the way.

David Swinbanks



This train is bound for the scrap yard.

Radio wars

WHEN television crews visit California's Mount Wilson observatory, their cameras sometimes fail inexplicably. The facility's own electronic equipment often behaves erratically. And radios are often plagued with interference. An encounter with alien radiation? Hardly.

The problem is right on the next mountain peak. There, a cluster of radio and television antennas pump out a total of over 10 million watts of broadcast programmes to the Los Angeles area — and subject any piece of unshielded electronic equipment in close range to the incidental indignities of electromagnetic disturbance. Researchers who use Mount Wilson's 100-inch reflecting telescope have been forced to become shielding experts as they swaddle their sensitive light-receiving equipment with enough metal to keep them working in a bath of radiation of the nearby antennas. Yet, despite their efforts, radiation from the antennas has essentially rendered large portions of the western sky off-limits.

Things could, however, be even worse. In 1989, the Sigma Communications Corporation proposed to set up another antenna farm on a second peak less than a mile away, this time to the south. With a potential output of another 10 million watts, the new facility threatened to shut down the observatory entirely, says director Robert Jastrow. After unsuccessful appeals to local authorities and the company itself, closing seemed a real possibility.

Then an observatory official happened to mention the problem to an attorney from the Los Angeles-based law firm Sheppard, Mullin, Richter & Hampton which agreed to take the case *pro bono*.

Now, after more than a year's negotiations, Sigma has given in. In an agreement completed this summer, the company agreed to restrict total electromagnetic in the direction of Mount Wilson to no more than current levels, to limit the height of the antennas, and to shield all lights at the antenna facility so that none are visible from the observatory.

The victory is only the latest for US observatories, which are becoming increasingly skillful in dealing with local governments and organizations to keep the skies free of radio and light interference. Palomar Observatory astronomers, plagued by radiation from a growing number of nearby low-power transmitters, have successfully lobbied the San Diego Board of Supervisors to change the zoning laws to include emission limits. And the University of California Lick observatory recently persuaded the nearby San Jose City Council to change most of the city's street lamps to low-pressure sodium designs that emit light only in the human visual range, where it can be more easily filtered out of observations. Christopher Anderson

Black gold causes a stir

London

A SWEDISH drilling company claims to have discovered crude oil in a granite rock formation — a location that should not contain oil, according to the currently accepted theory of oil formation, which holds that oil and natural gas are formed from biological material buried in sediments.

If the new finding is confirmed, it will lend some support to a controversial theory proposed by Thomas Gold from Cornell University. Gold argues that gas and oil deposits originate from methane, released from deep within the Earth, which has seeped up into sediments near the surface. He says he has no doubt that the new find supports his theory: "The case to me is completely clear."

The challenge to orthodox petroleum geology was met with scepticism last week by geologists and geochemists alike. Many said they would want more evidence that the oil could not have originated from sedimentary rocks in the surrounding area before they could accept the find at face value.

The Swedish claim was made on 8 October by Dala Djupgas Produktions, which is drilling deep into the granite rock below the centre of the Siljan Ring, an ancient meteorite impact site some 200 km northwest of Stockholm, in the hope of finding a large reservoir of natural gas. At 11:04 p.m. local time on 7 October, with the drill bit 2,833 m below the surface, oil began to wash up to the surface along with the usual mixture of rock debris and drilling fluid, says project manager Jack Kenney. Because no biological material should be present deep within a granite intrusion, the find would seem to support Gold's theory.

This latest claim to have struck oil is not the first to emerge from the project. Drilling on the present hole started only this year, but Dala Djupgas also claims to have pumped 80 barrels of oil in 1989 from a previous hole, drilled about 15 km away, near the crater's edge. That borehole, some 6,700 m deep, was abandoned last year, when the team ran into technical difficulties, Kenney says.

Dala Djupgas's first claim was dismissed by many petroleum geologists. Paul Philp, a University of Oklahoma geochemist who analysed samples of the 'black gunk' that Dala Djupgas extracted from the hole before pumping out the 80 barrels of oil, says that he could not distinguish between the samples from the hole and oil seeps found in shales near the surface in the Siljan area. Both, he says, contained similar 'biological marker' molecules—complex molecules derived from biological material (cholestane, a derivative of cholesterol, is an example).

The obvious explanation, says Philp,

was that oil had simply migrated down to the granite from sedimentary rocks near the surface.

Philp has declined to analyse any further samples from Siljan, arguing that his earlier findings were "taken out of context" by Gold and Dala Djupgas to support Gold's theory. But many geochemists suspect that the team was simply recovering the diesel oil drilling fluid that they had earlier pumped down the hole.

Kenney counters that extensive chemical analyses of the earlier samples show that they contain hydrocarbons other than those in the drilling fluid. And this time Kenney has a simpler argument to tackle the sceptics: the only drilling fluid that has so far been pumped down the second hole is, he says, fresh water.

"A lot of people are interested," says Chris Clayton from British Petroleum's Sunbury Research Centre in Middlesex. But he says that the industry will remain sceptical until Dala Djupgas provides samples for analysis. Kenney says that samples are being sent to a number of independent laboratories, but adds that the results will not be released until they have been reviewed by Dala Djupgas's eight-member scientific committee, which includes Gold. This may take several weeks, and even then, it seems unlikely that the samples will be made generally available, for commercial reasons.

Philp says that to be convinced that the new find does support Gold's ideas, he would want to see more geological data. If the rock surrounding the borehole is highly fractured (as it is likely to be, given the meteorite impact), he says, it is possible that the oil could have seeped into the granite from the surrounding shales. The Siljan Ring consists of a 40-km diameter intrusion of granite, surrounded by sediments.

The Swedish drilling project began in 1986, funded largely by the state-owned Swedish electricity utility Vattenfall, after Gold had convinced the organization that the Siljan site could hold huge reserves of gas or oil. Sweden has little coal, gas or oil, and with its population strongly averse to nuclear power, it is anxious to find new energy sources. But Vattenfall eventually withdrew from the project last year, which forced the drilling company to attract private funding to begin drilling the second hole.

The new find comes at an opportune time for the company, which is currently in the middle of a share issue to raise more capital for the project. But Kenney rejects any suggestion that the timing of the announcement will benefit Dala Djupgas. Under Swedish law, he says, the price of shares in a new issue is fixed before the shares are released. Although Dala

Djupgas shares may now increase in value during secondary trading, the company will receive no more money for each share it sells. Nevertheless, last week's news does seem to have increased interest in the Dala Djupgas share offer. "Everyone in Sweden wants to buy shares," observes company president Lars Bahlberg.

Peter Aldhous

VIDEO TELECONFERENCES

Long-distance hearing

Washington

LAST week's congressional hearing on the future of the US national laboratories looked — with one exception — just like any other, with its parade of witnesses appearing before the House committee on Science, Space and Technology, reading their prepared statements and responding to questions from the committee members. The difference was that those witnesses were thousands of kilometres from Washington.

The hearing was the first of a series scheduled for this congressional session that will use 'video teleconferencing' to save witnesses from flying in and out of Washington — a journey that costs West Coast residents a day and a half or two days of work to attend a three-hour hearing.

"I can envision the day when quick-turnaround hearings, multiple-point field hearings and regular interactions with witnesses testifying from abroad will all be done through video conference techniques," said committee chairman George Brown (Democrat, California) in his opening statement. "Getting to that point, however, will take some trial and error, such as what we will be experiencing this morning."

Actually, it was not as bad as Brown feared. Although the 15-frames-a-second technology (broadcast television uses 30 frames a second) made any sudden movements by the witnesses look jerky, that was a minor distraction. Audio problems blanked out a few words of testimony, but the sound was in general no worse than for most witnesses who appear in person. The least convenient part of the hearing was that only one channel was available, so that the committee had to hear first from one witness in Los Alamos, New Mexico, then switch to Livermore, California, to hear two more, and did not have the chance to go back and forth between the witnesses.

A number of federal agencies already use video teleconferencing to maintain contact among wide-flung outposts. The Department of Energy (DOE), for instance, uses it regularly to communicate with its national laboratories, and it was DOE's teleconferencing network that the committee borrowed for the hearing.

Robert Pool

From a sow's ear ...

Tokyo

JAPAN may have at last discovered what to do with its ill-fated nuclear-powered ship *Mutsu*, which has been sitting in port for most of the past 23 years. According to recent Japanese newspaper reports, the Science and Technology Agency plans to covert *Mutsu* into the world's largest oceanographic research vessel by ripping out its nuclear reactor and replacing it with a diesel engine.

Mutsu has a sad history. After launch in 1968, it took six years before local fishermen, who feared radioactive pollution, allowed the ship to set sail on its maiden voyage. No sooner had *Mutsu* put to sea than its reactor developed a leak. The crew desperately stuffed socks and rice into the reactor to try to stem the flow of radiation while the government offered the local fishermen in *Mutsu* City large sums in compensation to allow the ship back into port. They eventually agreed, but only on condition the ship leave afterward.

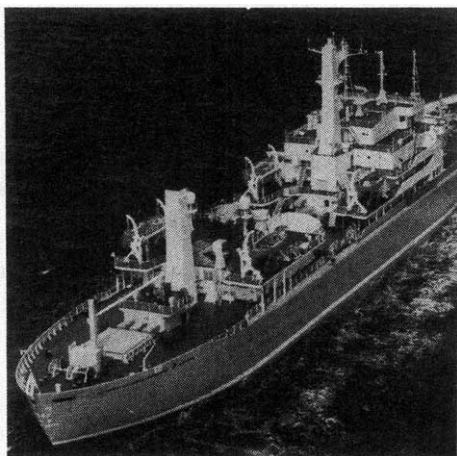
Mutsu was then towed from port to port in search of a home. But the ship was welcome nowhere, and after 14 years *Mutsu* returned to a custom-built port near its home city, where its presence is toler-

ated by suitably compensated fishermen. Last year, the Japan Atomic Research Institute, which runs the ship, cranked up the reactor (now encased in thick concrete) and took the ship on a few test runs in the ocean. But with no future plans to develop nuclear-powered ships, the government is in a quandary as to what to do with *Mutsu*, which has cost about ¥120,000 million (\$930 million) so far.

Thus the plan to covert *Mutsu* into a diesel-powered oceanographic vessel. Japanese newspapers say the Science and Technology Agency has calculated that *Mutsu* can be converted to diesel power for less than half the ¥10,000 million it would cost to build a new ship of the same size (past expenditures on *Mutsu* excluded). And at 8,242 tons, the refurbished *Mutsu* (suitably renamed to erase memories of the past) will be by far the biggest oceanographic research vessel in the world.

Agency officials refuse to confirm the reports. They say conversion to diesel power is just one option under consideration and no final decision has been made. Meanwhile, *Mutsu* waits patiently in port.

David Swinbanks



The *Mutsu* undergoing its first nuclear power test

INDIA

Saving female babies

New Delhi

USE of prenatal diagnostic (PND) techniques to determine the sex of a fetus has been prohibited under legislation introduced in the Indian parliament. It also bans advertisements of such tests and threatens violators with a fine of Rs10,000 and a three-year jail term.

The legislation is aimed at controlling the abortion of female fetuses. In a society where boys are preferred to girls, a growing number of private clinics have been set up to determine the sex of a fetus and help pregnant women undergo selective abortion of females. According to unofficial sources, some 2,300 such abortions take place every year in Bombay alone, and these clinics have begun to spring up even in semi-urban towns.

Under the new law, PND tests will be permitted only on medical advice for the purpose of determining genetic diseases, congenital anomalies, chromosomal ab-

normalities, and so on, but not fetal sex. Laboratories offering the tests must register with an appropriate authority. No doctor can conduct the tests on a pregnant woman unless she is either above the age of 35, has a history of spontaneous abortions, has been exposed to radiation or teratogenic agents or has a family history of mental retardation. Violation of any of the provisions of the law will land the doctor, as well as his patient, in jail.

To ensure that PND techniques are regulated according to the new law, the government plans to appoint a 14-member supervisory board headed by the minister in charge of family welfare. Each state will have a special committee to grant or suspend registration of laboratories, investigate complaints and enforce standards. All existing clinics that offer PND tests have been asked to register within six months or close down. It is likely that many will close down.

K. S. Jayaraman

Lessons from Japan

London

BRITAIN'S effort in high temperature superconductivity research and development may be dwarfed by the huge investment in equipment and manpower in Japan's industrial laboratories, but the prospects for a UK superconducting device manufacturing industry need not be bleak. British researchers still excel in certain niche fields that may be commercial hotbeds in years to come, and other niches are not yet full, according to a group of leading UK superconductivity researchers who have just returned from a government-sponsored fact-finding mission to Japan.

Gordon Donaldson, who led the mission, says that Japanese industry has a great "strength in depth" in superconductivity research.

But Donaldson says that the leading British research groups still compare very favourably with the best Japan has to offer, although he is concerned that too few British companies are working on applications for high temperature superconductors. He also warns that if Britain is to compete in commercializing high temperature superconductors, companies need to make better use of the country's major asset in the field: the handful of world-class superconductivity research groups in British universities and government laboratories.

The result is that the entire UK effort to commercialize the materials is extremely vulnerable, depending on the success of just a few companies, he says.

But Alford hopes to convince British industrialists that there are still profits to be made in high temperature superconductivity — provided that companies focus on those areas where British researchers excel, and work closely with academic research groups.

Donaldson singles out work on superconducting quantum interference devices, and microwave communications equipment using superconducting components, as British research strengths that hold commercial promise. (Alford's company already collaborates with groups at the University of Birmingham and the University of Manchester Institute of Science and Technology, in developing radio-frequency microwave communications equipment, but few other British companies have followed this example.)

The team also spotted one field where British industry has a chance to grab a market. Several Japanese companies are already working on silver clad high temperature superconducting wires, but there are fears that these products may not be made available to foreign companies. When the world develops a taste for high-temperature products like small superconducting magnets, high-quality superconducting wire may be in hot demand outside of Japan. Peter Aldhous

Agricultural research pruned

Heidelberg, Germany

GERMANY is to make dramatic cuts in agricultural research, an area that was of top priority in the former East Germany, but which holds much less interest to a newly unified country that is putting its resources into high-technology.

If last week's recommendations of the Wissenschaftsrat, Germany's science advisory council, are carried out, up to 50 per cent of the 4,000 scientists and 6,000 technical assistants working at the Academy of Agricultural Research and the research institutions formerly associated with the East German Ministry of Food and Agriculture will lose their jobs by the end of the year.

There is an surfeit in agricultural research because the former East Germany was striving for self-sufficiency, say Werner Kleinhanss and Susanne Reichrath from the Wissenschaftsrat. As the country had little foreign exchange to buy items from the West, it was forced to become self-contained in such areas as the development of scientific instruments and the production of chemicals. Since unification, that motivation has become less important.

The breeding and keeping of livestock and research on animal diseases were also of vital importance in East Germany. This led to the establishment of oversized institutions, such as the institute for research on livestock production at Dummerstorf near Rostock, which had 1,200 employees, including 350 scientists — three times the number at similar institutions in the west.

To set up competitive research for the future, the Wissenschaftsrat council suggests a major shift in emphasis in agricultural research, from self-sufficiency to ecologically oriented research — a move that will lead to severe cuts not only in personnel but also in research facilities. According to the science council, research areas of "outstanding scientific potential"

include animal and plant breeding as well as research on animal diseases and epidemics.

From the 47 institutes evaluated, the science council suggests founding five Blue List establishments (whose funding is shared equally by the Federal government and state in which the institute is located). They will be the first institutes of this kind in agricultural research in the united Germany. It also suggests that two national research laboratories should be founded: one in Jena in the state of Thuringia that will study animal epidemics, and one in Quedlinburg in the Harz for plant breeding. Five other institutes will be privatized, whose areas of expertise include machine building, veterinary medicine and food technology.

The biggest of the Blue List establishments will concentrate on soil research, ecology and environmental problems, especially those due to excessive use of fertilizer. Located at Müncheburg and Eberswalde in Brandenburg, it will employ 300 people (including 100 scientists) out of the 1,000 employees formerly employed there. The 1,200-person Dummerstorf institute will be trimmed to 70 scientists with up to 180 assistants, who will study genetics of livestock, breeding and feeding. The science council also suggests creating an institute in Berlin for Agricultural Development in Middle and Eastern Europe to do research aimed at easing the transition to a market economy.

The recommendations on agricultural research completes the process of reviewing 130 scientific establishments in the former East Germany. The Wissenschaftsrat has recommended creating more than 30 Blue List establishments, three national research laboratories and two Max Planck institutes. In all, a little more than 13,000 jobs will be available for the 30,000 employees of the former East German academics, which will cease to exist at the end of the year.

Barbara Bachtler

BRAZIL

War of the greens

São Paulo

JUST months before heads of state from around the world are due to converge on Brazil for the Rio 1992 environmental conference, Brazil's own environmental officials have become embroiled in a blood feud that has already claimed one job.

Tania Munhoz, who headed the Brazilian Institute of the Environment and Natural Renewable Resources, was fired earlier this year because, she said, she was no longer able to work with the explosive national secretary for the environment, José Lutzenberger. Since then, other offi-

cials, from the governor of the state of Amazonia to an Army general, have used the dispute as attack both Lutzenberger and Munhoz for compromising on Brazil's right to economic development and sovereignty over the rain forests.

The Brazilian environment, meanwhile, continue to suffer. Last month a team of scientists estimated that some 88,000 fires are burning in Brazil on a peak week, releasing between six and 12 million tons of soot — on the order of a volcanic eruption.

Ricardo Bonalume Neto

Bhopal case near end

New Delhi

AFTER nearly seven years of anguish, victims of the Bhopal gas disaster will at last be able to get their share of the \$470 million that the Indian government received two years ago as compensation from the Union Carbide Corporation of the United States. This follows a Supreme Court verdict last week upholding the compensation settlement reached in February 1989 by Union Carbide and the then-Congress government led by Rajiv Gandhi.

Although Union Carbide had paid the money in March 1989, its disbursement was halted by Gandhi's opponents and citizens' groups who described the deal as a sell-out to the US multinational. The successor National Front government referred the case back to the Supreme Court, in keeping with its election promise to get a better deal for the gas victims, while the money was allowed to lie in a bank gathering interest.

Reviewing the case afresh, four out of five judges of the Supreme Court declared the compensation settlement just and fair and asked the Indian government to make up for any shortfall. One judge dissented.

The majority verdict by the Apex Court has now paved the way for speedy disbursement of the money to nearly 200,000 survivors and families of the 3,800 killed when 40 tons of toxic methyl isocyanate (MIC) gas escaped from Union Carbide's pesticide factory in Bhopal on 2 December 1984. The court also asked Union Carbide to set up a 500-bed hospital in Bhopal and directed the Indian government to provide insurance coverage for unborn children whose mothers were exposed to MIC.

While upholding the compensation settlement, the Supreme Court reversed its 1989 decision granting Union Carbide and its Indian subsidiary immunity from criminal charges. Cases against Union Carbide and its officials pending in various Indian courts — which were quashed in 1989 — can now be reopened. In the original case filed in a Bhopal court, the former Chairman of Union Carbide, Warren Anderson, was charged with homicide. Several officials of the Indian subsidiary also faced criminal charges.

The Supreme Court verdict has generally been welcomed by Indians bar a few activists organizations which expected the court to raise the compensation amount. A spokesman of Union Carbide described as "unfortunate" the decision to allow criminal cases to be reopened. Union Carbide, which maintains that the gas leak was caused by employee sabotage, said that a "fair hearing on this sabotage issue will establish the true cause of the disaster."

With the Union Carbide lawyers insisting that Indian courts have no jurisdiction over the American company, it is not yet clear if the curtain on the Bhopal case has finally come down.

K.S. Jayaraman

The South answers back

Fernando Orrego

A recent UNFPA report on world population is full of errors, strongly biased and deceptive. The time has come to close down the agency and start again.

IN the context of a very bleak future, a recent report¹ by the United Nations Fund for Population Activities (UNFPA) talks of "the impending food crisis in many developing countries", concluding that "developing countries as a whole have suffered a serious decline in food self-sufficiency. Their cereal imports in 1969–71 were only 20 million tonnes. By 1983–85 they had risen to 69 million tonnes and are projected to a total 112 million tonnes by the end of the century." Yet the statistical tables in the same report show that in only three of the developing countries (those with a *per capita* income below US\$3,000) is there a decrease in *per capita* food production of 5 % or more in 1986–88, relative to 1979–81.

These countries (Botswana, Nicaragua and Guyana) represent 0.14 % of the world population. On the other hand, if one looks at food production in the 91 more-populated developing countries, containing 3,963.7 million people (97 % of the developing countries' population), weighted *per capita* food production rose in the same period (1986–88 versus 1979–81) by 27.2 %. By any standard, this is an excellent achievement, certainly no "food crisis". This is also in line with the long trend in world food production², as well as with the food glut of industrialized countries. Future prospects, such as the modernization of Soviet agriculture and the increasing application of biotechnology to agriculture, also look promising.

Education

According to UNFPA, "the low income countries, excluding India and China, have seen growth in primary school enrolments fall from an average annual rate of 5.6 per cent in 1975–1980 to 2.7 per cent in 1980–1987. According to UNESCO estimates, in 1985 approximately 105 million children 6–11 years old were not in school, and of those over 70 per cent were in the least developed nations. If current trends continue, by the year 2000, the number of out of school children will almost double to

approximately 200 million". But UNESCO's own figures³ show that primary-school gross enrolment ratios have increased in developing countries from 75.9 in 1960 to 99.8 in 1990. Regarding the rate of increase in enrolment, in 1975–60 it was 3.2 %, and in 1980–88 1.7 %. This decrease in rate is of course what happens when enrolment approaches saturation (a ratio of 100), as is seen in developed countries, where the 'increase' in 1975–80 was –0.2 %, and in 1980–88, 0.2 %. In developing countries, the rate of increase has been even higher in the second and third (higher) education levels (5.24 and 7.63 % mean annual rate, respectively, comparing 1988 to 1975).

UNFPA's report states: "The cost of providing education and health care from infancy to adulthood for a child in a developing country is much lower than in an industrialized country, but is still significant: about \$7,000 in India. By that measure, averting 106 million births since 1979, as India's official calculations show, represents a saving of \$742 billion. The uncounted costs — to the environment, to development prospects generally — are much higher". This looks like a promising procedure for increasing wealth in developing countries. But the figures are incorrect. According to UNESCO³, the annual expense in education per student in India was US\$55.7 in 1987 (latest available figure), whereas in health care, according to United Nations⁴ (also in 1987) the combined government and private expenditure was of \$8.65 *per capita*. Using the crude approach of multiplying the cost of health expense by 20 years, and that of education by 14, one reaches a figure of only \$952.80 for the combined expense. The proper way to obtain this figure is to calculate the net present value, which gives about \$232 for education and \$86 for health (using an interest rate of 10 % a year), or \$318 combined.

The assumption made by UNFPA for saying that "averting" a child's birth represents a saving of \$7000 only in education and health care, not to mention the improvement in development prospects, is that the child will be economically unproductive during his life. This can hardly be expected in a country whose gross domestic product has increased⁴ at an annual rate of 5.2% in 1980–88, and its population at 2.1%

annually^{1,3}. Using conservative estimates, each Indian male inhabitant makes a net economic contribution of about \$700 during his lifetime. Thus, by averting those births, instead of saving the sum calculated by UNFPA, India will suffer a net loss of \$74.2 billion.

Population densities

Perhaps the most fundamental demographic fact is population density, without which population growth rates (PGRs) are virtually meaningless. UNFPA's report wholly ignores the density factor, while stressing growth rates. This, in my opinion, leads to a distorted impression of developing countries as overcrowded and suffering from a population 'explosion'. Actually, the population density (inhabitants per km² of agricultural land) is: Africa, 80 (PGR 3%); Latin America, 58.2 (1.9%); Asia, 422.9 (1.8%); Europe, 213 (0.2%); North America, 55 (0.7%); Oceania, 15 (1.4%); Soviet Union, 69 (0.7%). This shows that virtually all the world's 'South' has a low density, indicating that there is no common population pattern for developing countries.

The world has other demographic problems, which UNFPA largely ignores. A common belief, at least outside Britain, is that life there is highly civilized, even graceful. The public image of China is quite different. But the population density of China, if also measured per km² of agricultural land, is 273, while that of the United Kingdom is 315. Mean densities, measured in relation to land surface, are China 115, and United Kingdom 234. This certainly indicates no more grounds for imposing the cruel policy of one child per family in China than in Britain. Such a policy, because of the preference for males, may lead in very short time to a demographic catastrophe.

The facts presented here demonstrate, in my opinion, that UNFPA's report is strongly biased and deceptive, as well as technically incompetent and substandard. As population issues in many senses are of great importance, and UNFPA seems incapable of analysing them, I suggest that the agency should be shut down and the responsibility for the task reassigned elsewhere. □

Fernando Orrego is in the Faculty of Medicine, Universidad de los Andes, Avda Ejercito 412, Santiago, Chile.

1. UNFPA *State of World Population 1991* (UNFPA, New York, 1991).
2. World Bank *World Development Report 1990* (Oxford University Press, New York, 1990).
3. UNESCO *Statistical Yearbook 1990* (UNESCO, Paris, 1990).
4. United Nations *National Accounts Statistics: Main Aggregates and Detailed Tables 1988 Part 1* (United Nations, New York, 1990).

A plea from Croatia

SIR — Your argument that “the best hope for Yugoslavia” is something other than “six sovereign republics” (*Nature* **350**, 176; 1991) has been frustrated by events. The majority of people in Slovenia, Croatia, Macedonia and Kosovo have voted freely for independence. The Federal Army has tried savagely to discipline Kosovo, Slovenia and Croatia, also threatening Bosnia, Herzegovina and Macedonia. The option of Yugoslav federation is definitely dead.

Yugoslavia was the nineteenth-century vision of prosperity for a group of states historically wedged between the dying Turkish Empire on one side and antagonism within the Austro-Hungarian Empire on the other. Between 1918 and 1990, the idea proved a failure and certainly cannot be considered a solution for our problems on the eve of the twenty-first century. But that does not mean that inhabitants of these countries should kill each other and destroy what they have accomplished in nearly three-quarters of a century and in more than a thousand years of separate existence. The brutality of the Federal Army, protecting Serbian territorial interests, has resulted in hundreds of Croatian deaths, more than a third of them civilians, as well as the destruction of towns and villages and particularly churches, hospitals and schools. Vukovar, Osijek, Pakrac, Zadar, Šibenik, Gospić, Varaždin — some of the larger towns — have been bombed, while Dalj and Kijevo, like many other communities, have been levelled to the ground by heavy artillery. The airports in Croatia are closed, harbours and ports have been blockaded by the navy. We cannot travel freely either abroad or within the country.

How can we talk science when we are afraid for our lives, in shelters? Can we plan on ordering chemicals and experimental animals when children are being killed and libraries burned? We would not want to share the fate of Mr Riadh Dinha Francis and his family (*Nature* **351**, 262; 1991), who were recognized on television by colleagues abroad and rescued from a Kurdish refugee camp by behind-the-scenes efforts. Nor do we want to be part of perhaps another DM1 million programme to assist researchers forced to leave their native land (*Nature* **351**, 261; 1991).

Science has no homeland, but scientists do, and they feel strongly about their countries. Nikola Tesla (1858–1943), Lavoslav Ruzicka (1887–1976) and Vladimir Prelog (born 1906) are three Croatian scientists known worldwide, who made their careers abroad, but always considered Croatia their homeland.

As scientists and citizens of the world,
NATURE · VOL 353 · 17 OCTOBER 1991

we despise frontiers and chauvinism of any kind. Why should we not collaborate with colleagues from Macedonia or Serbia? For scientific collaboration, it is not even necessary to live within the same country; to be forced to do that may even be counterproductive.

We believe that by alerting your public institutions and governments, as you have done in the case of fights for human rights in the Soviet Union, Chile, South Africa, Poland and Turkey, you could help us to stop the bloodshed and the destruction of our country and to establish the peace and economy indispensable for scientific work. Full recognition of Croatia as an independent and sovereign state is therefore urgent.

Our yearning for freedom is in line with that of every mature and self-confident person and nation. As a civilized European nation, with 1,300 years of cultural tradition, we Croats have every right to build our future by ourselves — after the war. Our expert studies have shown that Croatia can fulfil the economic and other needs of all its citizens.

The term ‘Balkanization’ has been used in the world’s media, including *Nature*. But we in Croatia strongly resent being called Balkans. This is not only geographical pedantry, but a result of the despicable flavour the word has acquired in the international vocabulary. ‘The Balkans’ has meant a barrel of dynamite, and the place where the First World War started, but it is well known who was responsible for all the Balkan wars, the present war against Croatia included. But why should we not, with international help, remake the relationships in the Balkans into the type of relationships existing in the Scandinavian countries?

Briefly, Croats are a nation and Croatia a state, and we are not Balkans in the sense in which you used the word. We have the right to be free, to live in democracy and peace and to make our living the way we judge best. It is our misfortune that Serbia, our neighbour, cannot escape the mediaeval spirit of attacking its neighbours’ territories. We may be mistaken in asking Europe and the rest of the world to understand the problem, but we do have the right to defend our homes and lives and the future of our children. At the same time, we hope we will survive as scientists and, some day, when our country is free and has recovered from the devastation, we shall still be able to publish in *Nature*.

IVICA VUCAK
ANTONIO JURETIC
Zagreb, Croatia

OSI procedures

SIR — In her letter (*Nature* **353**, 204; 1991) refuting allegations made earlier by Barbara Culliton concerning the procedures of the Office of Scientific Integrity (OSI), Dr Suzanne Hadley is very selective in her quotations from the *Federal Register*. She states that subjects of an OSI inquiry or investigation “are provided with an opportunity to review and comment on significant investigatory documents” but leaves out the remainder of that sentence: “. . . identified by OSI unless such disclosure would violate individual confidentiality or significantly impair the investigation”. Indeed, in the same paragraph, there is this statement: “No opportunity is provided for subjects to confront and cross-examine other witnesses interviewed by OSI.” Culliton is correct; these examples indicate why many scientists are aghast at the ‘star-chamber’ procedures of the OSI and at the feeble attempts to disclaim them.

PHILIP SIEKEVITZ
Rockefeller University,
1230 York Avenue,
New York, New York 10021-6399, USA

Human factors

SIR — In your article “Secret Service as ultimate referee” (*Nature* **350**, 553; 1991), you discussed the problem of apparent preference for particular digits in data recording. In addition to reports in the chemistry literature cited by Alan F. Thomas (*Nature* **352**, 9; 1991), this phenomenon has been recognized in several sciences. In a classic paper by G. Udny Yule (*J. R. statist. Soc.* **90**, 570–587; 1927) and a more recent review by D. A. Preece (*Statistician* **30**, 31–60; 1981), examples are included from agriculture, anthropometry, demography, geography, meteorology and physics.

Where investigators are spontaneous, they may show preferences for ‘round’ numbers or particular odd or even numbers, while different objective rounding rules are recommended by different authorities or built in electronically into calculators or other devices. Biases may also arise in conversion of units or other arithmetic manipulations. As Preece points out, this variety limits the usefulness of chi-square or other formal significance testing of departures from a uniform distribution. While this is likely to be the natural null hypothesis where the possible range of values is much greater than the resolution of measurement, the most appropriate alternative hypothesis will usually have been guessed in each case after detailed scrutiny of the data.

NICHOLAS J. COX
Department of Geography,
University of Durham,
Durham DH1 3LE, UK

Teaching the teachers

SIR — Mary Jane Drummond (*Nature* 352, 369; 1991) refers to the questionable scientific rigour of the “floating and sinking” test for seven-year olds, but there are even more fundamental problems in the teaching of science in primary schools through the new British national curriculum. The most basic is that the whole child population from the age of 5 to 11 is to be taught science by some 200,000 teachers, few of whom have themselves had any education in the subject. The primary teacher therefore needs deeper understanding of the basis of science than those who teach older children.

It is hard to see how teachers who do not themselves understand a concept, or who do not know how a conceptual hierarchy is built up, can possibly inspire children to develop such an understanding. This is especially so when a concept is counter-intuitive. We expect children to understand something about forces, to know that a push or a pull is needed to change the motion of anything. That

seems simple enough and to tie in with experience. But, the argument continues, if there is no push or pull, the motion is unchanged and the object goes on for ever. This is counter-intuitive. Everyone knows that just to keep a broken-down car moving steadily requires a hefty push. Stop pushing and the car stops. And in a way the intuitive approach is the correct one, for it is virtually impossible to escape from forces, and Newton's law is a hypothetical one. The likelihood that teachers who do not understand physics will innocently mislead or confuse their pupils and perhaps block the subsequent development of more advanced concepts must be considerable.

Could this be why Britain is a nation of non-scientists? Why so many people find science daunting and difficult? And if that should be the case, might we do better to leave primary schoolchildren without a science curriculum until such time as the teachers can be taught? Or should we put alongside the primary class teacher a specialist science teacher, trained to meet the challenge of introducing science to young children?

TOM MARGERISON
(Head of Education)

*British Nuclear Forum,
22 Buckingham Gate,
London SW1E 6LB, UK*

Shroud of Turin

SIR — Pierre Busson laments the various problems arising from the radiocarbon dating of the Shroud of Turin (*Nature* 352, 187; 1991). As I see it, there is only one problem: the dating itself. The Shroud of Turin is a religious object and, as such, it should never have been subjected to scientific scrutiny.

When Bill Libby developed the radiocarbon dating method, he had a steadfast rule: never to accept for dating anything related to religion. Radiocarbon dating has now shown that the Shroud was woven in the fourteenth century. So what? Scientists will say that that proves that the Shroud is a fake. Believers will say that it is an even greater miracle — the wrapping of the body of Christ in a shroud that was woven 1,300 years later.

Religion is perfect and unchangeable, the work of God. Science is imperfect, everchanging, and, I suspect, the work of the Devil. The two should never be mixed. The scientists who participated in the dating of the Shroud of Turin should repent and promise never to do anything like that again. Creationists are even more guilty, for they have been mixing science and religion for years and years. They should abandon their evil practices forthwith, lest the wrath of God descend upon them like a ton of bricks.

CESARE EMILIANI

*Department of Geological Sciences,
University of Miami,
Coral Gables, Florida 33124, USA*

Genome ethics

SIR — John Maddox's update on the Human Genome Project (*Nature* 352, 11–14; 1991) deals fleetingly with the ethical problems posed by this research. The author mentions the concern that individuals endowed with wrong genes may be subjected to insurance discrimination, and tends to play down German sensitivity regarding eugenics, believing such applications unlikely. However, there are reasons for concern.

After the visually discernible chromosomal irregularities are investigated, the genetic cartographers will tackle the genes that define our anatomical and physiological features, because these, given our long biological history, very probably represent the bulk of our genetic makeup.

Inevitably, the investigators will chance upon those genes that have originated late in our development and which define the anatomical and physiological substrates of human cognition. Our very consciousness may be found to be gene-dependent. This will have serious consequences.

Our manipulation of the genetic constitution of animals and plants for economic reasons is quite extensive, even though our knowledge of their genome is

limited.

Is there any doubt, once it becomes clear what underlies human intelligence and memory, that this knowledge will be put to use in the name of individual and national interest? Can we afford not to engage in eugenics? Human-directed evolution is inevitable and as it is faster than natural evolution it will be within its power to change our race within a few generations.

Who will have access to this knowledge and when, how and to what end? As most if not all of the funding of the project is public, the public should determine these matters.

EMIL ASCHHEIM

*12716 Headwater Circle,
West Palm Beach, Florida 33414, USA*

Wittgenstein

SIR — J. R. Smythies has tried¹ to resolve the long discussion of Wittgenstein's merits by diagnosing him as insane. In particular, he mentions paranoia, schizoid personality disorder and schizophrenia and also cites some reported life events and the “schizophrenese” speech disorder, for which neither a scientific definition nor a citation is given. We cannot agree with these psychiatric statements.

(1) The diagnosis of schizophrenia, and, even more, that of schizoid personality disorder, is a very difficult and uncertain procedure even in the assessment of living patients. But the diagnosis does not rely on a patient's ‘schizophrenese’ writing². Moreover, Wittgenstein's life argues against a diagnosis of schizophrenia. In retrospect, a diagnosis of cyclothymia seems much more likely. Also the family history and experiences of living relatives of Wittgenstein, which cannot be discussed in detail for ethical reasons, argue against schizophrenia.

(2) The description of ‘schizophrenese’ writing applies not only to Wittgenstein's work, but to that of many other philosophers and theologians^{3,4}.

(3) Even a diagnosis of psychiatric illness does not of itself diminish the content of the thoughts and ideas of particular patients. Psychiatry should try to get the better of the common prejudice against insanity.

P. FISCHER
B. KÜFFERLE
P. BERNER

*Psychiatrische Universitätsklinik,
Universität Wien,
A-1090 Wien, Austria*

1. Smythies, J. R. *Nature* 350, 9 (1991).
2. *Diagnostic and Statistical Manual of Mental Disorders*, American Psychiatric Association 2nd edn (Washington, DC, 1987).
3. Rochester, S. R. & Martin, J. R. *Crazy Talk: A Study of the Discourse of Schizophrenic Speakers*. (Plenum, New York, 1979).
4. Schwartz, S. *Behavioural and Brain Sciences* 5, 579–626 (1982).

Celebration of a dead genius

The second centenary of Michael Faraday's birth is being celebrated in Britain with great style, but in a way that inevitably provokes regret as well as pride.

"Now I know what's wrong with British science", said a sharp-tongued mid-Atlantic person at the end of last Friday evening's glittering occasion at the Royal Institution of London: "It's the cult of the amateur. People in dinner-jackets doing magic tricks for other people in dinner jackets." The complainant had forgotten that, until recently, the performers had been required to wear white ties and the tail-coats that go with them. (Then, as now, there were hardly ever women.) And, in any case, last Friday's show (for *that* it was) was a nostalgic evocation of times past by people now alarmed about the future.

The occasion was part of a veritable season of celebration of the second centenary of the birth of Michael Faraday, on 22 September 1791. (The celebrations will go on at least until Faraday's portrait is replaced on the verso of the £20 British banknote.) The men in black ties were the two immediate past directors of the Royal Institution — Lord Porter (president of the Royal Society until a year ago) and Sir John Meurig Thomas, who was succeeded in the summer by Dr Peter Day and who now divides his time between structural chemistry and running what many consider to be the unrunnable — Britain's University of Wales.

The setting is unique, indeed unbeatable, so to speak. In down-town London, two hundred yards across Piccadilly from the Ritz Hotel and almost immediately opposite Brown's Hotel, there is a colonnade of fake Grecian columns (put up in 1838) concealing the London-brick structure in which Count Rumford (otherwise Benjamin Thompson, who had fled the United States after being accused of spying) founded in 1799 a centre for the popularization of the "mechanical arts".

The fake colonnade is a perpetual headache for the managers of the place, who are often advised that, unless they spend sums of money they believe they cannot afford, it may come loose from the structure it adorns. That encloses a lecture-room of artless but unsurpassed design — half a hemisphere, of radius is no more than six or seven metres, in which more than 400 people (when the uncomfortable balcony is full) can hear the man in the black tie at the centre even if he whispers. Humphry Davy (who succeeded Rumford) evidently had in mind an anatomy theatre in which the exhibits would not be living bodies or cadavers but demonstrations of

the mechanical arts.

Showmanship is the essence of the Royal Institution, which is thereby true to Rumford and to Davy (who told a friend of his own "weak, glorious, pitiful, sublime conceited egotism"). Even now, the set-piece lectures, called discourses, are held on Friday evenings, are timed to begin exactly at 9 p.m. and to finish exactly 59 minutes and 50 seconds later. Demonstrations (and the applause they prompt, especially when they end in loud bangs) play a crucial part, but take literally incalculable seconds. If speakers cannot *ad lib* (or abbreviate in mid-peroration), the striking of the clock may be surreptitiously delayed.

Last Friday's practised performers required no such artifice. One (Porter) started bang on time and the other (Thomas) finished just on the witching hour. Between them, they managed more than one demonstration every three minutes.

Porter, acting as if relieved that the constraints of being president of the Royal Society have been sloughed off, delivered the more elegant half-discourse. His jokes were mostly half-sentences, not even one-liners. His barbs were distributed to politicians, administrators and researchers in equal measure. At one point, he sat inside an electrically conducting cage charged by an electrostatic machine to show, with the help of an electroscope and an insulated wand carrying coloured tufts of wool, that the electric charge of a conducting body lies on the surface — as Faraday had done at the same spot in 1836.

Thomas, no less accomplished, has a different style. Controlled passion, called *hwyl* by people of his Welsh cultural persuasion, restrains side-swipes at innocent bystanders and is best fitted to solemn occasions — funerals certainly, birthday celebrations (even 200 years on) perhaps. Yet this was the half-discourse that reminded the audience that Faraday was first a chemist and only afterwards the founder of electromagnetism.

Thomas's most evocative demonstration, tenuously linked with Faraday's celebrated lectures on "The Chemical History of the Candle", had an assistant (also in black tie) pouring a bucketful of carbon dioxide evaporated from dry ice into a container on one side of a crude balance arm to show that the density of carbon dioxide is greater than that of air. Two weeks earlier, as part of the same celebrations, Sir John Cadogan (research director

of British Petroleum) had in the same spot used gas chromatography to show that the benzene isolated and characterized by Faraday in 1825 was 99.7 per cent pure.

That Faraday was a genius, and an Englishman to boot, is not disputed. Welsh Thomas came close to English heresy last Friday in comparing him to Newton, but he will be forgiven. What the glittering occasion also showed is that, for the Royal Institution, Faraday is a domestic genius whose presence is still felt. For half a century after he was hired by Davy in 1812, he lived and worked at the Royal Institution, leaving his meticulous notebooks for the scholars who still pore over them. Last Friday, they were showing off these volumes as well as his notes taken at four of Davy's lectures, bound by Faraday himself, with which this untutored person had introduced himself to his employer.

It was, in short, a great and sentimental occasion, to which people had been invited jointly by the Royal Society and the Royal Institution. How could it have evoked even a scintilla of sourness?

The black-tie part is irrelevant. In Davy's mind, the Royal Institution was a means of making propaganda for the mechanical arts among the well-to-do of Mayfair and thus of raising money for research. Over the years, the Davy-Faraday laboratories occupying the less grand upper rooms have indeed sheltered people such as John Tyndall (one of the most prolific contributors to *Nature* after its foundation in 1869) as well as, more recently, J.D. Bernal, Kathleen Lonsdale and Max Perutz. Faraday himself said that the need to perfect experiments for the weekly discourses spurred his scientific work. As the French are now well aware, there is no harm, but a power of good and even public profit, in banging the drum for research among the rich and influential.

It matters more that Faraday's heroic tale is in danger of becoming hackneyed in the retelling. True, he was a blacksmith's son, but the modern middle classes had hardly been invented then. And there is a danger that dwelling on his undisputed innocence of mathematics may recreate the impression that the mechanical arts are, or should be, simple after all. Ministers in last Friday's audience may even have left with the impression reinforced that British science would not be complaining about its budget if its practitioners were half as smart as Faraday.

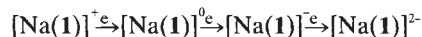
John Maddox

A new challenge to valency

Edwin C. Constable

A FUNDAMENTAL tenet of chemistry is that isolable compounds should be electrically neutral overall — it is not possible to keep a bottle of sodium ions on the shelf. Over the past few years a number of compounds have been described that apparently conflict with this view, and these have provided a challenge to theoreticians and synthetic chemists alike. Lehn and his co-workers have now described another example of such a compound, which has been described as a 'molecular element'¹.

Lehn and colleagues are interested in the properties of metal ions 'imprisoned' within cage-like molecules² such as the cryptand, compound **1** in Fig. 1. The reaction of **1** with sodium bromide yields the complex [Na(1)]Br in which the sodium cation is trapped within the cage and is bonded to the six nitrogen atoms of the 2,2'-bipyridine subunits. The passage of an electric current through solutions of [Na(1)]Br results in current flow at three different potentials corresponding to the formal processes:



The compound [Na(1)] is formed at the electrode as blue-violet crystals, which are immediately destroyed upon exposure to the air. The crystal structure of this compound has been determined and it shows only an [Na(1)] unit (Fig. 2) — no anions appear to be present.

The compound could contain a neutral sodium atom within the cage, or it could

now well-established⁴, and are associated with encapsulating ligands such as **1**. However, in these electride salts the electron is located in relatively large 'empty' sites within the crystal lattice where one might expect an anion. Unfortunately, no such vacancies are found within the lattice of the blue compound, in which the [Na(1)] units are tightly packed. Where else could the electron be?

Lehn and colleagues propose that their compound represents a compromise between the charge-separated structure of an electride, [Na(1)]⁺e⁻, and that of an encapsulated metal atom, [Na⁰(1)]. They suggest that the unique properties of the 2,2'-bipyridine units in ligand **1** allow the electron to reside on the pyridine rings of the cage. In other words, the structure may be represented as [Na⁺(1⁻)], in which (1⁻) represents the radical anion of the ligand. Evidence of this comes from the displacement of the sodium from the centre of the cage towards one of the three 2,2'-bipyridine units (Na-N_{short} ≈ 2.6 Å; Na-N_{long} > 2.8 Å). One of the properties of 2,2'-bipyridine which makes it of such value to the transition-metal coordination chemist is the presence of low-lying unoccupied π* orbitals which provide a 'sink' for any excess electron density on metal centres⁵. It is suggested that the 2,2'-bipyridine residue that the sodium is closest to has accepted that missing electron, and now possesses the character of the bipy⁻ radical anion. The authors propose the use of the generic term 'cryptatium' for these compounds, with [Na⁺(1⁻)] being sodio-cryptatium.

This now offers the intriguing possibility for the synthetic chemist to fill the continuum of structural types between a neutral metal atom and the spatially separated electrides. If the cage ligand does not contain

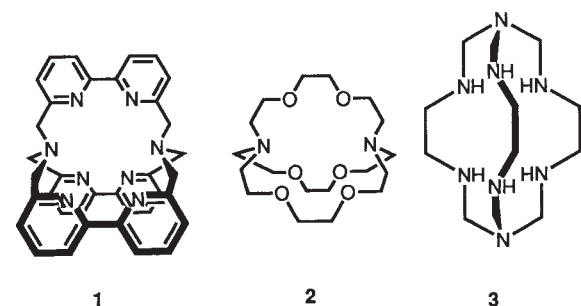


FIG. 1 Three macrocyclic 'cryptand' ligands.

contain a positive sodium ion with the 'missing' electron residing elsewhere. A sodium ion is much smaller than a sodium atom, and the crystal structure clearly indicates that it is an ion which is present (indeed the structure is very similar to a related cryptand complex [Na(crypt)]Br which unambiguously contains Na⁺)³. Is the blue material an isolated [Na(1)]⁺ cation, or is the elusive electron present?

One possibility is that the anion required for electrical neutrality is simply an electron. Such 'electride' salts are

a 2,2'-bipyridine, then the charge-separated species are obtained. For example, the reaction of the cryptand **2** (Fig. 1) with one equivalent of potassium metal gives shiny black crystals of the electride salt, [K(2)]⁺e⁻ which behaves like a molecular metal⁶. However, the situation is not quite so simple, and even in these cases there are other sites in which the electron may reside. Of these, one of the less likely is on another metal ion to give a metal anion, but this is exactly what happens when **2** is allowed to react with an excess of sodium to give

beautiful gold-coloured crystals of the salt [Na(2)]⁺Na⁻, which contains a 'sodide' anion⁷. This structure challenged many of the then existing conceptions regarding the alkali metals.

The unusual properties conferred upon metal ions when they are trapped in a cage have been explored by many workers. For several years, Sargeson has made detailed studies of the properties of metal ions within cages such as **3** (Fig. 1), and has demonstrated the remarkable stabilization of 'unusual' ox-

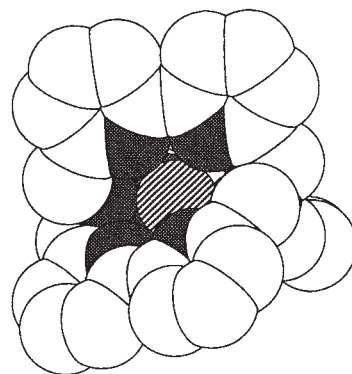


FIG. 2 Space-filling model of the compound [Na(1)]. Open circles, carbon; shaded, nitrogen; striped, sodium.

idation states upon encapsulation⁸. For example, the reduction of [Co(3)]³⁺ to [Co(3)]²⁺ is facile and gives a cation that is very long-lived in solution. This is in contrast to the properties of conventional cobalt(II) amine complexes, which are very labile and would be expected to dissociate. The stability is ascribed to the inability of the cobalt(II) ion to 'squeeze' between the bars of the cage in which it is imprisoned.

The ambiguities associated with the formalization of charge distributions are well-known to inorganic and bio-inorganic chemists, and there are many molecules such as 2,2'-bipyridine whose behaviour may be described as non-innocent. Dioxygen is a prime example of such a ligand, and there is still controversy over the electronic nature of the oxygenated form of haemoglobin⁹. Should the Fe(O₂) pair be formulated Fe^{II}(O₂), Fe^{III}(O₂⁻) or Fe^{IV}(O₂²⁻)? Complexes of 2,2'-bipyridine have traditionally presented problems, and many apparently low-valent compounds should be formulated⁵ as ligand and radical

1. Echegoyen, L., DeCian, A., Fischer, J. & Lehn, J.-M. *Angew. Chem. Int. Ed. Engl.* **30**, 838-840 (1991).
2. Lindoy, L. F. *The Chemistry of Macrocyclic Complexes* (Cambridge University Press, 1989).
3. Caron, A. et al. *Helv. Chim. Acta*, **68**, 1577-1582 (1985).
4. Dye, J. L. & Huang, R.-H. *Chem. Brit.* **26**, 239-244 (1990).
5. Constable, E. C. *Adv. Inorg. Chem.* **34**, 1-63 (1989).
6. Huang, R.-H. et al. *Nature* **331**, 599-601 (1988).
7. Tehan, F. J., Barnett, B. L. & Dye, J. L. *J. Am. Chem. Soc.* **96**, 7203-7208 (1974).
8. Sargeson, A. M. *Chem. Brit.* **15**, 23-27 (1979).
9. Hughes, M. N. in *Comprehensive Coordination Chemistry*, Vol. 6 (eds Wilkinson, G., Gillard, R. D. & McCleverty, J. A.) 541-754 (Pergamon, Oxford, 1987).

complexes such as $[M^{III}(\text{bipy})_3]$ rather than $[M^0(\text{bipy})_3]$. The niceties of this ambiguity are of direct relevance, as the complex cation $[\text{Ru}(\text{bipy})_3]^{2+}$ plays a vital role in many chemical photoconversion systems. The critical excited state $[\text{Ru}(\text{bipy})_3]^{2+*}$ is best described¹⁰ as charge-separated $[\text{Ru}^{III}(\text{bipy})_2(\text{bipy}^-)]^{2+*}$.

The new compound described by Lehn and co-workers, provides a structural model for this photoexcited state as well as challenging our preconceptions about the ionic state. These new compounds

may be regarded as 'expanded' elements, in which the effective size of the alkali metal atom has been increased — instead of finding the electron in one of the valence orbitals of the metal ion it is found slightly out in the orbitals of the directly bonded ligands. □

Edwin C. Constable is in the Cambridge Centre for Molecular Recognition, University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

MANTLE GEOCHEMISTRY

Diamonds deliver the dirt

J. J. Gurney

As nearly impenetrable and inert capsules, diamonds carry within them tiny mineral inclusions trapped as they grew. On page 649 of this issue¹, Eldridge *et al.* use ion microprobe analyses of the sulphur and lead isotope ratios in sulphide inclusions to learn something of the geochemistry of the mantle rocks in which they formed. In particular, their results emphasize that some diamonds contain sediments subducted, with oceanic lithosphere, from the Earth's surface.

From the mineralogy of diamond inclusions, it has already become established that diamonds come in two classes — peridotitic and eclogitic — reflecting the mantle source rocks in which they formed². It is also becoming apparent³ that peridotitic diamonds are derived from primary mantle carbon with a limited range in $^{13}\text{C}/^{12}\text{C}$ ratio and that these diamonds may be Archaean in age (over 2,500 million years old)⁴. Conversely, there is considerable evidence for the formation of younger eclogitic diamonds with widely variable carbon isotope ratios interpreted to have originated in recycled (subducted) crusts.

Unfortunately these observations are not conclusive. The equivocal nature of the evidence permits several schools of thought, amongst them one which holds that eclogitic diamonds too are formed from primary carbon, but in an isotopically inhomogeneous mantle⁶. Eclogite — broadly the high-pressure equivalent of basalt or gabbro — would be a product of partial melting of the mantle in this model. The carbon isotopic heterogeneity would have to be largely inherited from the source, as high-temperature fractionation processes do not seem capable of creating it. A parallel wide variation in oxygen isotope ratios occurs in the silicate minerals in eclogite rocks⁷ entrained in kimberlite (extinct volcanic pipe) in which diamonds are found. Oxygen isotope fractionation can be strongly developed only

at low temperatures, so that mantle heterogeneity or crustal recycling are again two counterpoised interpretations.

The new data of Eldridge *et al.* show that sulphides associated with the same eclogitic diamond class have sulphur isotope-ratio variations well beyond those expected in the generally accepted primary mantle. The cumulative evidence that three stable isotopes C, S and O show signs of pronounced fractionation in eclogitic minerals is far more persuasive than any of these results in isolation. All these variations can be readily achieved at low temperatures. Geochemists faced with a choice between radically reassessing the range in isotope ratios for these elements within the primary mantle and accepting a recycled crustal origin for diamond eclogites, must surely now opt for the latter.

An additional piece of the jigsaw, previously missing, comes with the lead-isotope evidence in the new work, which indicates that the recycled component includes continental sediments. If eclogite is recycled oceanic lithosphere, some sediments must also accompany the subducted crustal volcanic material down to regions of high temperature and pressure. A missing component that is, volumetrically, even more significant remains to be defined: where is the extensive oceanic peridotite that must have been the major component of the down-going slab?

This slab, essentially the oceanic lithosphere at the time, would consist of surface and near-surface sediments and igneous components, plus the underlying volumetrically dominant depleted mantle peridotite from which the oceanic crust formed by partial melting. According to known ages of eclogitic diamonds, these subduction events occurred in the Proterozoic⁸. The older Archaean ages and mantle carbon signatures have been found in garnets from garnet harzburgites⁴ — rocks so depleted in their major elements that they have the

RÉSUMÉ

Early bird

How a routine re-examination of a modest fossil reptile turned into a search for the closest relatives of birds is recounted by A. R. Milner and S. E. Evans in *Palaeontology* (34, 503–513; 1991). Fragmentary material of *Lisboasaurus estesi* from Guimarota in Portugal, about 160 million years old, was originally described as from a lizard. Further study by R. Estes revealed the presence of socketed teeth, more characteristic of archosaurs (crocodiles, dinosaurs and birds). Milner and Evans now show that *Lisboasaurus* most closely resembles the troodontids, a group of advanced, birdlike theropod dinosaurs usually found in younger sediments. More tentatively, some aspects of the dentition place the fossil within the Avialae, a clade that includes *Archaeopteryx* and all other birds. *Archaeopteryx* dates to about 150 million years ago, so the existence of other, more-or-less contemporaneous birdlike reptiles seems reasonable, say the authors.

On the scent

THE nose, twitching in the wind, will catch a fugitive fragrance on a receptor in a cilium, embedded in the mucus layer of the olfactory epithelium. The receptor (one of many) is linked to a G-protein, which in turn control an ion channel activated by cyclic AMP. Efficient capture of the molecule that represents the odour evidently depends on the mucus matrix, and this, it turns out, contains an abundant protein called olfactomedin (D. A. Snyder *et al.* *Biochemistry* 30, 9143–9153; 1991). The polypeptide chain has a relative molecular mass of 57,000 and exists as a disulphide-linked dimer, heftily glycosylated. There are loads of this material in specialized cells, in which it is presumably made, as well as in the olfactory epithelium. But just what it is doing remains to be seen.

Flash and burn

THERE'S no smoke without lightning, or not where large forest fires are concerned. Don Latham (*J. geophys. Res.* 96, 17,151–17,157; 1991) describes a fire, started deliberately for forest management one dry summer afternoon in Ontario, that consumed 460 hectares in 3 hours. Within half an hour of the start, a plume of smoke had risen 6–7 kilometres into the sky which was drawn out into an anvil shape by gentle wind shear. Latham recorded 37 cloud-to-ground lightning strikes, most of them composed of positive charge. In thunderstorms, such strikes are not only rare but also dangerous, because they have the highest currents. But Latham suspects the plume resembles more those emitted from volcanoes, which carry net positive charge and have also been seen to cause lightning.

most magnesian olivines in the mantle and contain no clinopyroxene. These harzburgites are thought to represent the metasomatized residues from an unusually large degree of partial melting or from successive cycles of partial melting in the mantle more than 3.2 thousand million years ago.

As both peridotitic and eclogitic diamonds are much older than the great majority of the relatively young volcanic rocks (less than 200 million years old) that ultimately sample them and transport them to the surface, they have to be stored in the cool, thick, chemically reducing, brittle lithospheric keel characterized by the roots of the old continental nuclei^{9,10} — they would not survive in the deeper, hotter, more oxidizing and convective asthenosphere. Apparently, the younger eclogite is underplated onto the ancient continental core which contains the harzburgites⁵.

Subducted oceanic lithosphere may well be represented by compositions more iron-rich than these very refractory rocks. The only obvious component linked to diamond genesis that could fit this description is clinopyroxene-bearing garnet lherzolite. Garnet lherzolite xenoliths found in kimberlite sometimes contain diamonds¹¹ and garnet lherzolite minerals are found as inclusions in diamonds from a wide geographical spread and in large numbers of kimberlites and lamproites¹². At present it is not clear whether these diamonds associated with garnet lherzolite formed at the same time as Archaean harzburgitic diamonds, as eclogitic diamonds, or at some other time.

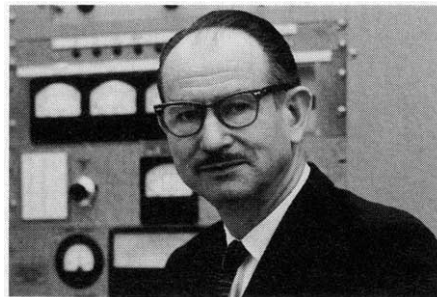
The association of diamonds with garnet lherzolite should be further investigated to fill in an important gap in this rapidly unravelling story. Given the difficulty in obtaining suitable diamond inclusions for most analytical approaches, the SHRIMP ion microprobe used by Eldridge *et al.* is the ideal instrument with which to tackle this question. □

J. J. Gurney is in the Department of Geochemistry, University of Cape Town, Rondebosch 770, South Africa.

Edwin M. McMillan (1907–1991)

EDWIN Mattison McMillan, a leading light in the golden age of nuclear physics and a key innovator in the quest for powerful 'atom smashers', died on 7 September at the age of 83 of complications from diabetes, seven years after a debilitating stroke. Apart from the war years, McMillan's scientific career was spent at the University of California at Berkeley, in particular in the laboratory founded by E. O. Lawrence.

McMillan is renowned for the discovery (with P. H. Abelson) of the first



McMillan by the controls of the Bevatron.

transuranic element, neptunium and the discovery (with G. T. Seaborg, J. W. Kennedy and A. C. Watis) of the second, plutonium. He shared the 1951 Nobel prize in chemistry with Seaborg. He is also famous for the independent discovery of the principle of phase stability for particle accelerators, for which he and V. I. Veksler were awarded the 1963 Atoms for Peace award. This principle, along with strong focusing, is crucial to the realization of the microscopes of present-day particle physics — the very large accelerators at CERN, Fermilab and other institutions.

McMillan was born in southern California on 18 September 1907. Educated in science at the California Institute of Technology, he did his graduate work under E. U. Condon at Princeton University, receiving his PhD in 1932. Coming to Berkeley on a National Research Fellowship, he soon joined Lawrence's group and shortly after was appointed to the Berkeley physics faculty. Among his early accomplishments were the quantitative study of electron-positron pair production by γ -rays in the Coulomb field of an atomic nucleus in 1934 — at a time when quantum electrodynamics was in its infancy and in need of verification — and the discovery of oxygen-15 the same year with M. S. Livingston.

In 1939–40, during a study of the fission of uranium, McMillan found indications, then (together with Abelson), proof of the creation of the first transuranic element ($Z = 93$), named neptunium by them. In the latter part of 1940, he began a search for the next transuranic element using the powerful 60-inch cyclotron, but had to interrupt his efforts to

undertake work on radar (and, later, on the atomic bomb). With McMillan's blessing, Seaborg and colleagues pursued his line of attack and in early 1941 discovered an isotope of plutonium ($Z = 94$).

In the summer of 1945, McMillan discovered the important principle of phase stability for particle accelerators. He showed that charged particles orbiting and being accelerated periodically by a radiofrequency electric field organize themselves into stable bunches that can be raised to very high energies by slowly varying the frequency of the accelerating field and (if a constant orbit radius is desired) increasing the confining magnetic field. It was the breakthrough necessary to extend the maximum energy of electron and proton accelerators, first to hundreds of megaelectronvolts and ultimately to the giga- and teraelectronvolt range. McMillan immediately moved to implement his ideas by construction of a 330-MeV electron synchrotron. J. Steinberger and W. Panofsky used its high-energy photon beam in 1950 to discover the neutral π -meson.

The discovery of phase stability was, in fact, anticipated by Veksler, whose papers, published in the Soviet Union in wartime (1944), were unavailable to McMillan. In a short note in 1946, McMillan graciously acknowledged Veksler's priority.

Upon Lawrence's death in 1958, McMillan became director of the Radiation Laboratory (now the Lawrence Berkeley Laboratory), serving for 15 years before his retirement in 1973. He provided steady, sensible leadership to the Radiation Laboratory during its heyday as the world's finest high-energy physics laboratory, but was, in my opinion, miscast as an administrator. He was first and foremost a physicist's physicist. After retirement, he spent a happy year at CERN in 1974–75, with the group making a precision measurement of the magnetic moment of the muon. I recall his delight in telling me how, by detailed analysis of a tiny unexplained loss of the muons from their orbits, he had traced the cause to the very slight unevenness of the magnet pole faces and had even been able to deduce the depth and width of the microscopic grooves made by the milling machine.

McMillan was a man of quiet mien, modest, fair-minded and with firm principles. Respected and well-liked by colleagues, he was known to his family as a fun-loving parent, sometimes testing, sometimes teasing, always sharing his tremendous curiosity in the natural world.

J. D. JACKSON

J. D. Jackson is in the Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA.

- Eldridge, C. S., Compston, W., Williams, I. S., Harris, J. W. & Bristow, J. W. *Nature* **353**, 649–653 (1991).
- Meyer, H. O. A. in *Mantle Xenoliths* (eds Nixon, P. H. *et al.*) 501–522 (Wiley, Chichester, 1987).
- Sobolev, N. V. *et al.* *Dokl. Akad. Nauk USSR* **249**, 1217–1220 (1979).
- Heimstaedt, H. & Gurney, J. J. in *Kimberlites and Related Rocks* (ed. Kornprobst, M. C.) 425–434 (Elsevier, Amsterdam, 1984).
- Richardson, S. H., *et al.* *Nature* **310**, 198–202 (1984).
- Deines, P. *Abstr. 5th Int. Kimberlite Conf.* 74–75 (1991).
- Garlick, G. D., MacGregor, I. D. & Vogel, D. E. *Science* **172**, 1025–1027 (1971).
- Richardson, S. H., Erlank, A. J., Harris, J. W. & Hart, S. R. *Nature* **346**, 54–56 (1990).
- Boyd, F. & Gurney, J. J. *Science* **232**, 472–477 (1986).
- Haggerty, S. E. *Nature* **320**, 34–38 (1986).
- Davison, J. B. & Smith, J. *Nature* **254**, 580–581 (1975).
- Gurney, J. J. in *Kimberlites and Related Rocks* Vol. 2 (ed. Ross, J.) 735–965 (Blackwell Carlton, 1989).

Not an open-and-shut case

Stephen W. Jones

KINETIC models for the behaviour of ion channels can now be interpreted as testable hypotheses about channel structure and function. This exploits the structural information resulting from cloning of ion-channel proteins, and the rich kinetic detail available with patch-clamp recording. One relatively neglected topic, the recovery of channels from inactivation, is considered by Ruppersberg, Frank, Pongs and Stocker on page 657 of this issue¹, and by two other groups^{2,3}, who all find that channels often reopen on the way back from inactivated to closed states. This is at once expected and surprising: a reopening current is predicted by simple models of the inactivation process, but it has not been previously observed, and I think that few suspected that it really exists. Reopening currents have implications for the molecular mechanisms of channel gating, and they could affect the complex patterns of electrical activity characteristic of neurons and other excitable cells.

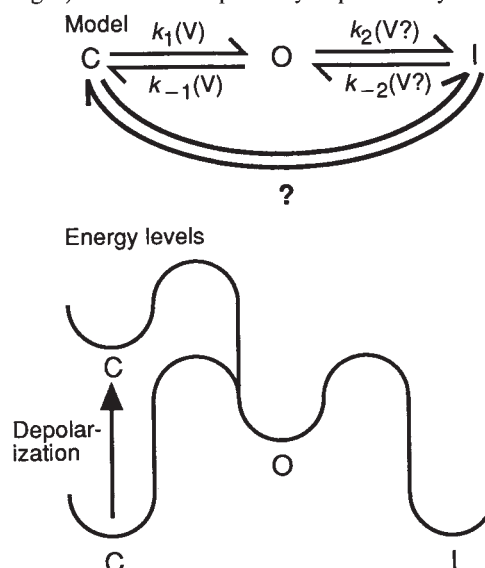
Many voltage-dependent ion channels open only transiently in response to a maintained depolarization. This suggests at least two kinetically distinct processes: the initial channel opening ('activation') and the subsequent transition to an 'inactivated' state. The simplest kinetic scheme for inactivation is a three-state model with a conducting open (O) state and two distinct nonconducting states: a closed state (C), which predominates at negative voltages, and the inactivated state (I) (see figure). Realistic models for the behaviour of ion channels tend to be much more complicated, with multiple closed states (and often multiple open and inactivated states), but the three-state model captures the essence of the problem.

At the macroscopic level (that is, in recording the activity of a large population of channels), inactivation is voltage-dependent. Stronger depolarization causes faster and more complete inactivation. But for a variety of channels there is good evidence that the microscopic inactivation process (O→I) is not voltage-dependent⁴⁻⁷. In this view, channels are pulled through the open state by mass action, and the apparent voltage-dependence of inactivation arises from kinetic coupling with the strongly voltage-dependent activation process.

One simple physical interpretation is the 'ball-and-chain' model⁴: inactivation results from physical plugging of the channel pore by a part of the protein molecule (the 'ball') that is linked to the rest of the protein by a flexible 'chain'. Binding of the 'ball' can also prevent the

channel from closing. Site-directed mutagenesis has provided direct evidence for this ball-and-chain model⁸, and synthetic 'ball' peptides block potassium channels in much the same way as in natural inactivation⁹.

At the macroscopic level, recovery from inactivation is also voltage-dependent (more rapid at negative voltages). This can be partially explained by



The three-state model for inactivation of voltage-dependent ion channels. In its simplest form, only activation (C→O) is directly affected by voltage, with channel opening being faster (and closing slower) at depolarized voltages. On the cyclic version of the model, closed channels can inactivate (usually slowly), and channels can return from the inactivated state without reopening. The energy-level diagram suggests how the linear model allows macroscopic inactivation to be voltage-dependent. At negative voltages, the closed state is the lowest energy level, so most channels are in that state. Upon depolarization, the closed state is destabilized, and channels open transiently before passing into the inactivated state, which is now the lowest available energy level, even though the relative positions of the open and inactivated states have not changed.

coupling to the activation process. However, for strong inactivation to occur, the I→O step must be slow, which limits the maximal rate of recovery. Such a limiting rate is sometimes observed⁷, but for at least some channels there appears to be an additional, intrinsic voltage dependence to recovery from inactivation¹⁰.

One counterintuitive consequence of the linear C→O→I model is that channels must pass through the open state on the way back from I to C. In principle, this would produce a current during recovery from inactivation as the channels transiently reopen. Such a reopening current

was not observed for sodium channels⁴ or T-type calcium channels⁷, which led to the conclusion that there must be a direct pathway from I→C, requiring (minimally) a cyclic three-state model. Physically, this would mean that the channel can close even while the channel mouth is blocked by the 'ball'. But a reopening current has now been reported for both potassium^{1,2} and calcium³ channels. Calcium channels may reopen during recovery, but most do not³. A large fraction of potassium channels², possibly all of them¹, recover via the open state. Ruppersberg *et al.*¹ also conclude that both inactivation and recovery are intrinsically voltage-dependent, as they find that a putative 'ball' peptide binds and is released in a voltage-dependent manner. Release (corresponding to recovery) is considerably faster at more negative voltages^{1,2}, possibly because entering K⁺ ions 'push' the 'ball' out of the channel². Binding of the 'ball' may also depend on K⁺ and voltage¹, but more weakly².

These observations may have implications for channel function. For calcium channels, a reopening current could provide a slow, steady trickle of calcium ions into the cytoplasm, with consequences for release of neurotransmitters and for other metabolic processes³. For potassium channels, a reopening current could produce a long-lasting afterhyperpolarization, as Ruppersberg *et al.* suggest. Their current-clamp experiment is clever: many will be surprised to hear that a patch has a resting potential. But the high (> 10 GΩ) input resistance of the patch would tend to overestimate the effect of the reopening current in a real cell. The significance of the reopening current will depend on the details. If channel closing is much faster than recovery from inactivation, reopenings will be brief and unsynchronized, so the reopening current will be small. The relatively large reopening currents observed by Ruppersberg *et al.* may

reflect the unusually slow closing rate (time constant about 120 ms) that they report for their channels. Thus, it re-

1. Ruppersberg, J. P., Frank, R., Pongs, O. & Stocker, M. *Nature* **353**, 657–660 (1991).
2. Demo, S. D. & Yellen, G. *Neuron* (in the press).
3. Slesinger, P. A. & Lansman, J. B. *Neuron* (in the press).
4. Armstrong, C. M. & Bezanilla, F. *J. gen. Physiol.* **70**, 567–590 (1977).
5. Cota, G. & Armstrong, C. M. *J. gen. Physiol.* **94**, 213–232 (1989).
6. Zagotta, W. N. & Aldrich, R. W. *J. gen. Physiol.* **95**, 29–60 (1990).
7. Chen, C. & Hess, P. *J. gen. Physiol.* **96**, 603–630 (1990).
8. Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Science* **250**, 533–538 (1990).
9. Zagotta, W. N., Hoshi, T. & Aldrich, R. W. *Science* **250**, 568–571 (1990).
10. Keynes, R. D. *Proc. R. Soc. Lond. B* **243**, 47 (1991).

mains to be established whether reopening currents are a general property of inactivating channels, and whether the reopening currents are of significant amplitude (especially at physiological concentrations of extracellular K^+). And as Ruppertsberg *et al.* note, most known afterhyperpolarizations result from Ca^{2+} -dependent potassium channels,

GALACTIC ASTRONOMY

A census of the Milky Way

Gerry Gilmore

WITHIN astronomical accuracy, there are about as many people on Earth as there are luminous stars in the Milky Way. It is, however, rather more difficult to make a reliable census of the latter than the former. The ethnographer can travel to obtain representative or complete samples and can ask specific questions of interest — age, place of birth and so on. The astronomer, while able to invade the privacy of nearby stars, has been limited to chance visits by travellers from afar for a glimpse of the larger scale picture, and can deduce only some of the fundamental parameters of interest from feasible observations. That limitation is however changing, as the combination of new technology and a revival of interest in the structure, origin and evolution of the Milky Way are providing detailed knowledge of the several stellar populations which make up our Galaxy. These results are presented in a variety of recent and forthcoming papers (refs 1–6 and C. Soubiran *et al.*, manuscript in preparation) which are providing a detailed description of the earliest stages of evolution of the Milky Way, and the properties and distribution of its oldest stars.

That knowledge depends on the fact that observers are now analysing in detail large samples of stars far from the neighbourhood of the Sun, thus minimizing the worst effects of galactic myopia. It also depends on the fortunate circumstance that astronomy is easier than archaeology, in that astronomical fossils are still alive for study. Any star ever formed with less than about three quarters of the mass of the Sun still shines and remains pretty well unchanged since its formation. Thus, by finding ever older stars, the astronomer can find direct measures of the evolution of the Milky Way. The distribution of stars by age maps when they formed, the distribution in space maps where they formed, the distribution over chemical element abundances maps the history of the creation of the chemical elements, the distribution in kinematics maps the rate at which the gas cloud which became the Milky Way cooled and radiated

which are different from the A-type potassium channels studied here. But reopening currents are still intriguing from both the practical and theoretical points of view. □

Stephen W. Jones is in the Department of Physiology and Biophysics, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA.

its energy, and so on.

The concept of 'stellar populations' is conveniently exploited, in part to allow some organization of this wealth of information, but largely because stars do seem to be describable as members of a small number of classes. The stars of the visible Milky Way, the galactic disk, are predominately young or middle-aged, are typically similar in chemical elemental abundances to the Sun, and have a lot of angular momentum but small random motions. By contrast the

evolution — probably involving the merging of a small number of identifiable lumps over a very long time — is favoured. Discrimination between these possibilities requires proof that the thick disk is, or is not, intermediate in all the important parameters defining a population: age, chemical abundance, orbital energy and spatial distribution.

The recent surveys agree in assigning kinematics to the thick-disk that are indeed intermediate between the halo and disk populations. The vertical velocity dispersion is nicely twice that of the old disk and half that of the halo. These numbers are not explicable by assuming that the thick-disk stars used to be part of the thin disk, and were ejected in some way. Rather, they favour formation from a distended gaseous distribution. This is supported by the age distribution, which shows stars in the thick disk to be predominately (and perhaps exclusively) older than the oldest in the thin disk.

There are two somewhat puzzling aspects of the results, the understanding of which will form the next big challenge in deducing the history of the Milky

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The Sombrero galaxy, similar to our Milky Way, has a large, metal-poor, hence blue, halo.

stars far from the galactic plane are exclusively old, are lower in heavy elements by a factor of typically 50 than the Sun, have little or no angular momentum but have large random motions.

A question of common interest to recent studies of the structure of our Galaxy is: what (if anything) lies between these extremes? A few years ago it became well established that a stellar population — the thick disk — lies between the main disk and halo populations, in that its spatial distribution was intermediate between them. It was (and is) unclear if the thick disk is also intermediate in an evolutionary sense. If it is truly intermediate, and is not simply the tail of either of the other populations with which it shares our part of the Galaxy, then apparently galactic formation was a relatively smooth and continual affair. If not, then a discontinuous

Way: the boundary between the thin disk and the thick disk is remarkably sharp in both kinematics and chemical elemental abundances, and the overlap between the thick disk and the halo is remarkably small. That is, the thick disk is almost too discrete a stellar population. There is quite a distinct break in the distribution of the kinematics (a measure of the amount of cooling of the pre-stellar gas cloud) and in the element ratios of those elements created exclusively in short-lived massive stars (a measure mostly of the rate of star formation up to that time) between the thick disk and the thin disk, at an abundance of about one-third that of the Sun.

In this case the apparently discrete nature of the thick disk really is remarkable, as the chemical elements formed by the earliest halo stars should have had the same small angular momentum as do

the present halo stars, and so should have fallen towards the Galactic Centre. Similarly, one does not expect a thing as large and complex as a galactic disk to form either rapidly or smoothly. A diversity of properties is expected. An intermediate stellar population as far from the Galactic Centre as we live should have a sizeable precursor population with a low chemical abundance, high angular momentum and intermediate velocity dispersion, as should the main galactic disk. It should not be as easy as it seems to be to distinguish the thick and the thin disks.

No doubt such complexities are real, and await the completion of an even more impressive stellar census for their elucidation. These recent surveys have

shown the way, revitalized study of the early evolution of the only major galaxy we are likely to map in detail, and made a substantial contribution to our understanding of the state of the part of the Universe which became the Milky Way. □

Gerry Gilmore is in the Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK.

1. Crosswell, K., Latham, D. W., Carney, B., Schuster, W. & Aguilar, L. *Astr. J.* **101**, 2078–2096 (1991).
2. Ryan, S. & Norris, J. *Astr. J.* **101**, 1835–1864; 1865–1878 (1991).
3. Rose, J. A. *Astr. J.* **101**, 937–949 (1991).
4. Rose, J. A. & Agostinho, R. *Astr. J.* **101**, 950–966 (1991).
5. Nissen, P.-E. in *Elements and the Cosmos* (ed. R. Terlevich) (Cambridge University Press, in the press).
6. Ratnatunga, K. U. & Yoss, K. *Astrophys. J.* **377**, 442–461 (1991).

ANTIGEN PRESENTATION

The second class story

Ronald N. Germain

THE CD4⁺ and CD8⁺ T lymphocytes take a very narrow view of the world around them. They recognize only cell-surface complexes of peptides bound to the highly polymorphic membrane proteins encoded by class I and class II genes of the major histocompatibility complex (MHC). In the past year there has been considerable progress in the analysis of peptides bound to class I molecules, and on page 622 of this issue Rudensky *et al.*¹ now report the first comparably precise characterization of physiologically produced peptides associated with class II proteins.

The technical approach used to analyse both the class I- and class II-bound peptides is deceptively simple, involving acid elution from immunoaffinity purified MHC molecules, followed by microsequencing. Application of this method to class I-derived peptides gave the first identification of naturally processed antigens^{2–5}, with rather surprising results. The products were invariably short octa- or nonapeptides that were 100–1,000 times as potent in T-cell assays as the longer synthetic analogues that had been in routine use, mostly due to markedly enhanced stable binding to the class I molecule^{6,7}. A second remarkable result was that the recovered peptides corresponding to a given T cell-recognized determinant had uniquely identifiable N and C termini. Finally, if multiple peptides associated with a single class I allelic product were sequenced, certain positions showed a strong bias for a particular amino acid ('motif residues')^{4,5}.

The data on class II-associated peptides differ in each of these regards. Of the 12 peptides characterized, all are 13 and some at least 17 residues in length.

Several overlapping peptides corresponding to the same core determinant were identified, having precise N termini but varying in length and hence in C-terminal residues. Peptides eluted from a single allelic class II product showed no positions with conserved residues. The proteins yielding the class II-binding peptides were also distinct in their cellular location from those giving rise to the class I binders. Peptides from the class I molecule HLA-B27 mainly derived from intracellular/cytosolic proteins such as histones and stress proteins⁵; the class II peptides were derived from membrane glycoproteins or serum proteins known to enter the acidic vacuolar compartment in large quantity. These latter results provide support for a model put forward in these pages nearly five years ago⁸, suggesting that class I and class II proteins differ primarily in their capacities for acquiring peptides from these two distinct protein pools.

Short peptides

Three observations provide a potential explanation for why only short peptides are associated with class I molecules. First, peptide plays a critical role in driving stable post-synthetic assembly of the polymorphic class I heavy chain with β -2 microglobulin^{9–11}. Without suitable peptides available, such dimer formation is inefficient, class I transport to the cell surface is impeded, and the few class I complexes that do reach the surface rapidly denature at physiological temperature¹². The effectiveness of a peptide in mediating this stabilization of class I structure is highly dependent on both its sequence and length^{6,7}. The optimal octa- and nonapeptides naturally associated with class I are best at pro-

moting class I heavy chain folding and stable association with β -2 microglobulin, primarily by decreasing the dissociation rate of the assembled complex⁷.

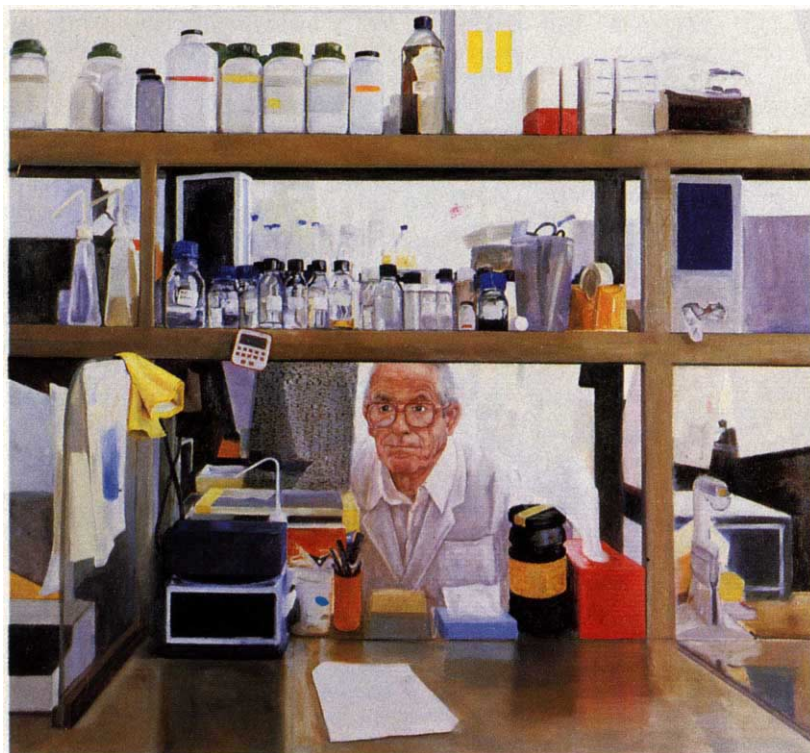
Second, peptides corresponding to those recoverable from class I molecules are not found free within the cell and, conversely, only cells with class I molecules able to bind the peptide in question have a detectable content of that peptide¹³. The most likely explanation is that the peptide pool feeding class I binding sites rapidly turns over and also probably consists of peptides larger than those found tightly associated with class I. Binding to class I seems to protect the peptide from complete degradation, but does so efficiently only when the peptide optimally fits the binding site. Third, the direct crystallographic imaging of nonameric peptide in the HLA-B27 class I molecule¹⁴ reveals a combination of motif-residue side-chain packing into polymorphic binding pockets and bonding of the N and C termini of the peptide to conserved residues of class I, such that longer peptides would perforce lack important stabilizing bonds and/or interfere with proper class I folding. The buried location of the peptide bonds of the terminal residues would serve to protect them from peptidases, but only after trimming the peptide to a length that allowed burial to be achieved.

Different structures

Given the predicted similarities of the class I and class II peptide binding sites, and the ability of the same synthetic peptides to bind both classes of molecules, why then are the naturally bound ligands so different in structure? The use of distinct processing enzymes in cytosol and acidic vesicles, and active peptide selection by the permease responsible for importing peptides into the endoplasmic reticulum for class I binding, could be part of the explanation. Peptidases that mediate trimming events at the site of peptide-MHC interaction could also vary in specificity and activity. In addition, even a few differences between the class II and the class I binding sites in such features as salt bridges and conserved residues that act to anchor peptides could significantly change the ligand length optimum for the two molecules, as Rudensky *et al.*¹ speculate.

However, the fundamentally different roles of peptide in controlling overall MHC molecule structure and transport are more likely to be major contributors to the differences seen. Effector CD8⁺ T cells need to identify those cells intrinsically producing the protein precursor of the peptide to be recognized and to avoid destruction of innocent bystanders. This necessitates a mechanism for class I loading that assures early stable

Portrait of a biochemist



With the hanging of the portrait shown here, Frederick Sanger joins other British luminaries of contemporary science in the National Portrait Gallery in London. Appropriately, the picture was put on view earlier this week, just before the recipient of the 1991 Nobel prize for chemistry was announced. Sanger has won the award twice: in 1958 (outright) and 1980 (shared). □

occupancy with peptide, and minimizes emergence from the endoplasmic reticulum of class I readily able to acquire peptide from extrinsic proteins. Linking stable class I assembly and transport to peptide binding accomplishes this goal. The contribution of peptide to maintenance of class I structure needs to be large, otherwise too many 'empty' molecules might escape the endoplasmic reticulum. This limits useful peptides to those with optimal fit, as others do not make a sufficient contribution to stabilization.

The CD4⁺ T cells, especially those interacting with B lymphocytes, are not similarly focused on proteins having a unique endogenous association with the presenting cell. Rather, class II deals with peptides derived from proteins largely captured from without. These peptides are produced in an acidic compartment, and studies *in vitro* show that class II acquires peptide best when it is offered ligand at low pH (4–6), then is brought to neutrality as would occur during passage to the cell surface^{15–17}. The class II dimer must thus itself be able to resist dissociation in this denaturing environment. In fact, substantial amounts of class II are lost during this passage through the processing compartment¹⁸ and it is likely that, once free of protection afforded by invariant

chain, class II must rapidly acquire peptide and leave the processing compartment or, as shown here¹ for Ex itself, become fodder for the peptide-generating apparatus.

In contrast to class I, neither initial class II dimer assembly, nor intracellular transport into or out of the peptide-loading compartment seems dependent on tight peptide binding¹⁸. Instead, class II appears to interact with peptide only with low affinity (fast-on/fast-off) in the processing compartment. Such loose peptide–class II complexes then probably leave the acidic processing compartment, at which time class II undergoes a conformational change which itself stabilizes a closed or compact structure of the binding site, and traps the peptide that previously was loosely associated. Although peptide appears to dictate whether class II follows the folding path to this stable structure, the peptide itself does not contribute nearly as much to the maintenance of the folded state as it does for class I molecules¹⁹. Thus, templated trimming to give unique termini is not necessary before stabilization and site closure can be attained. This model predicts the capture of a variety of related peptides containing only enough appropriate structure to give low-affinity binding in the acidic processing compartment, consistent with the ragged ends

seen by Rudensky *et al.*¹. Optimal occupancy of a polymorphic pocket may also not be required for adequate pre-closure binding, accounting for the absence of an easily identified motif in the sequences being reported.

Biological implications

The biological implications of the different rules governing the stringency of fit for class I and class II binding peptides are substantial. The requirement for optimal fit based on motifs for occupying polymorphic pockets and proper spacing of these residues from anchor residues within the length constraints of the class I site severely limits the number of suitable peptides generated from any given protein. This can account for the small number of functional determinants recognized by CD8⁺ T cells in association with any given class I allele that derive from even a large set of viral proteins²⁰. But a suitable CD8⁺ target cell need only present one suitable peptide species from among those that can be made from all viral gene products. Such stringency however places marked limitations on the likelihood of effective CD8⁺ T-cell surveillance for mutant self-proteins. Conversely, class II would be much less selective and, as observed, find numerous useful determinants in even a single protein. B cells capturing individual proteins via surface immunoglobulin could thus present a suitable determinant to CD4⁺ T cells in many cases.

Numbers

A second major point made by Rudensky *et al.* is quantitative rather than qualitative. They note that more than 50 per cent of the eluted material from a given population of class II molecules consists of just 5–7 distinct peptides. Based on estimates of the minimum number of peptide–MHC complexes needed to trigger a T cell^{21,22}, they suggest that cells in heterozygous individuals might be constrained to effective class II presentation per allele of only a limited (fewer than 100) distinct species of self-peptides. This would of course be only a small sampling of the possible peptides derived from the estimated 10,000 different proteins each cell synthesizes, and thus much of the complement of self-peptides would be invisible to the developing and mature T-cell pool, with important implications for self-tolerance and repertoire generation. Do these numbers hold up and are they different from either the expected ones or those relevant for class I?

First, some real numbers. Macrophages, dendritic cells and activated B cells have perhaps one to several hundred thousand class II molecules on their surface, not the 20,000 cited by Ruden-

sky *et al.*¹ for small B cells. Assuming that the 5–7 predominant peptides represent 50 per cent of all bound peptide (an estimate that is likely to be too high, as only the best binding peptides are preserved during class II molecule purification), this would leave perhaps 100,000 class II molecules for occupancy by other peptides. With 100–200 complexes per cell as the number for minimal T-cell stimulation, this would allow for 500–1,000 different species of peptide–MHC complexes to serve as self-tolerogens, thymic selectors or allogeneic targets for each allelic class II product expressed. Obviously the distribution of ligand concentration will not be so even. However, the T-cell hybridoma assays used to estimate minimal effective ligand density are at least ten times less sensitive than assays with cloned T cells, making it likely that, for some T cells, only 10–20 complexes per peptide species are needed for activation. Thus, even with a large fraction of class II occupied with only a limited peptide set, the universe of potential receptor ligands is still large. Furthermore, these numbers are not very different from those suggested for class I. The measured number of sites occupied by an efficiently recognized and naturally processed viral peptide is 200–500 per cell²³. With several hundred thousand class I molecules per cell, this again suggests a ligand repertoire of about a thousand per allelic class I

product.

It is thus less the total number of peptide–MHC species that is remarkable about the present results. Rather, it is the nature of the highly represented peptides and what they say about how class II gets filled that really matters. All those that could be identified derive from molecules either known (E α , invariant chain, BSA) or likely to enter and be present in the endocytically related processing compartment at very high local concentration. At least one peptide, that derived from E α is associated with the compact form of class II dimer reported to identify the cohort of molecules with tightly bound peptide¹⁷. As only a fraction of all newly synthesized class II dimers enters this characteristic peptide-associated conformation¹⁸, this suggests a relative paucity of peptide in the processing/loading compartment. Antigen capture by the processing cell is thus a key limiting feature of the class II-presenting system. The practical import is that perhaps more attention should be paid to delivering adequate amounts of antigen to class II antigen-presenting cells than to finding optimum peptides, whereas the opposite might be true for

class I.

Finally a few words about a second paper from Rudensky *et al.*, on page 660 of this issue²⁴, in which the authors identify the determinant seen by monoclonal antibody Y-Ae (ref. 25) as a complex of processed E α peptide (residues 56–73) and A α^b A β^b . This antibody is thus an analogue of a peptide-specific, MHC-restricted T-cell receptor. The remarkable feature of this antibody is that it readily identifies its ligand on medullary cells of the thymus of animals expressing both A α^b A β^b and E α , but fails to do so on cortical epithelial cells that are nevertheless positive for A α^b A β^b expression²⁵. It remains to be seen why the E α –A α^b A β^b complex is not formed by these latter class II-expressing cells, but as the authors write in a paraphrase of a classic understatement²⁶ “The implications of these data for positive selection. . . and negative selection. . . have not escaped our attention”. Nor will they escape the attention of the immunological community. □

Ronald N. Germain is in the Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892, USA.

EVOLUTION

Hypercycles spring to life

Robert M. May

A CENTRAL problem of early ‘prebiotic’ molecules that replicate themselves is how information is accumulated and handed on from generation to generation. And how can complex structures evolve accurately if unsupervised by some sort of correction mechanism? The hypercycle hypothesis¹ gets around this problem by proposing a cyclic sequence of fairly simple molecules, each catalysing the replication of the next and eventually building up a structure of greater net complexity. This idea ran into some problems early on², but has now unexpectedly received new life from a report in *Physica* from Boerlijst and Hogeweg³.

Eigen and Schuster were the first to emphasize that, in the absence of error-correcting mechanisms, the length of molecules in a self-replicating system is restricted by the accuracy of replication¹. At the mutation rates found in prebiotic conditions, the maximum possible string-length appears to be short; that is, larger molecules on average cannot be accurately replicated. Thus there is an information threshold posing a barrier to the evolution of more complex and information-rich molecules^{1,4}. The hypercycle theory is essentially a device to cross this information threshold, so even though each molecule in the

hypercycle is still constrained to the relatively short maximum possible string-length, we can still effectively build complex structures.

In a characteristically lucid and incisive review of these ideas, Maynard Smith raised an apparently devastating objection: hypercycles are vulnerable to ‘parasites’². Maynard Smith defines a parasitic molecule (number 4 in the cycle, say) as one whose replication gets catalytic support from the precursor molecule in the hypercycle (number 3 in this instance), but does not give catalytic support to any other molecule. Clearly, if this parasitic molecular species 4 gets more support from its precursor species 3 than does the regular hypercyclic species 4, then the parasite will be selected and the hypercycle chain broken. Thus hypercycles would appear to be evolutionarily unstable, unless explicit appeal is made to group selection.

But now Boerlijst and Hogeweg claim that “spatial self-structuring in incompletely mixed media can profoundly change the outcome of evolutionary processes”³. In particular, they show that in such spatially structured situations hypercycles can easily purge themselves of the kinds of parasites envisaged by Maynard Smith. They consider a

1. Rudensky, A., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. & Janeway, C. A., Jr *Nature* **353**, 622–627 (1991).
2. Van Bleek, G. M. & Nathenson, S. G. *Nature* **348**, 213–216 (1990).
3. Röttsche, O. *et al.* *Nature* **348**, 252–254 (1990).
4. Falk, K., Röttsche, O. & Rammensee, H.-G. *Nature* **348**, 248–251 (1990).
5. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **353**, 326–329 (1991).
6. Schumacher, T. N. M. *et al.* *Nature* **350**, 703–706 (1991).
7. Elliott, T., Cerundolo, V., Elvin, J. & Townsend, A. *Nature* **351**, 402–406 (1991).
8. Germain, R. N. *Nature* **322**, 687–689 (1986).
9. Townsend, A. *et al.* *Nature* **340**, 443–448 (1989).
10. Lie, W.-R. *et al.* *Nature* **344**, 439–441 (1990).
11. Townsend, A. *et al.* *Cell* **62**, 285–295 (1990).
12. Ljunggren, H.-G. *et al.* *Nature* **346**, 476–480 (1990).
13. Falk, K., Röttsche, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290–296 (1991).
14. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
15. Jensen, P. E. *J. exp. Med.* **171**, 1779–1784 (1990).
16. Harding, C. V., Roof, R. W., Allen, P. M. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7170–7174 (1991).
17. Sadegh-Nasseri, S. & Germain, R. N. *Nature* **353**, 167–170 (1991).
18. Germain, R. N. & Hedrix, L. R. *Nature* **353**, 134–139 (1991).
19. Dornmair, K., Rothenhäusler, B. & McConnell, H. M. *Cold Spring Harb. Symp. quant. Biol.* **54**, 409–416 (1989).
20. Bennick, J. R. & Yewdell, J. W. *J. exp. Med.* **168**, 1935–1939 (1988).
21. Harding, C. V. & Unanue, E. R. *Nature* **346**, 574–576 (1990).
22. Demotz, S., Grey, H. M. & Sette, A. *Science* **249**, 1028–1030 (1990).
23. Falk, K. *et al.* *J. exp. Med.* **174**, 425–434 (1991).
24. Rudensky, A., Rath, S., Preston-Hurlburt, P., Murphy, D. B. & Janeway, C. A., Jr *Nature* **353**, 660–662 (1991).
25. Murphy, D. B. *et al.* *Nature* **338**, 765–768 (1989).
26. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737–738 (1953).

chain of hypercyclic interactions taking place in an inhomogeneous medium that is represented by a square matrix of 300 times 300 (or 90,000) cells. Each cell, in any one generation, contains either a molecule of a particular species or is empty. The molecular chemistry of the hypercycle works itself out through processes of decay (the cell becomes empty), replication (a molecule in a cell can replicate into an empty neighbouring cell) and catalysis (the probability of such replication is increased if there are catalytic molecules in at least one adjacent cell). Movement or diffusion among neighbouring cells operates in between each generational 'time step', according to rules that are standard for such cellular automaton models^{5,6}.

In a manner that is becoming increasingly familiar (since first studied in models of the Belousov-Zhabotinski reaction)⁵⁻⁹, Boerlijst and Hogeweg find that their model gives spiral waves that continually move outward from foci that themselves shift over time, giving rise to rich patterns as spiral waves move and collide. Boerlijst and Hogeweg also give an insightful account of the biology of what is going on. Within the spiral, there is a very unequal distribution of darwinian fitness, with the molecules in the middle of the spiral generating the eventual offspring of the entire spiral (in cyclic sequences of molecular species), whereas the molecules at the edge of the spiral disappear. That is, the spatial self-structuring gives rise to a small sub-class of molecules which dominates reproduction.

When a parasitic molecular species of the Maynard Smith variety appears in this spatially structured world, its fate is very different from the selfish success it would enjoy — to the eventual destruction of the entire system — in a homogeneous setting. Instead, Boerlijst and Hogeweg show in detail that such parasitic populations necessarily grow in the direction of the outward movement of the spirals, and thereby are destroyed. Thus parasites appear, and successfully grow, but always toward the outer periphery of the spirals, where all molecules are extinguished. This conclusion is robust, deriving from the spatial structure of the system: Boerlijst and Hogeweg show that even parasites with very large replicative advantages are readily purged soon after they appear. The only exceptions arise if a parasite originates exactly at the centre of a spiral; in this event the parasite can survive as a relatively small 'cyst', encapsulated at the spiral's centre (and if it ever escapes, then it is swept to the spiral's edge and removed).

Ever since the rich and varied patterns generated by John Horton Conway's Game of Life (based on a cellular auto-

maton) were brought to the attention of a wide audience by Martin Gardner's column in *Scientific American*, many scientists have been fascinated by cellular automata as metaphors for living things. As surveyed by Langton¹⁰ and others (mostly with backgrounds in physics), a growing literature is based on the view that the self-structured patterns seen in cellular automaton models may help us understand prebiotic replicators, and thence early evolutionary processes. These studies of cellular automata are a lot of fun, but I do not believe they are likely to tell us much about the origins of life, any more than do ordinary crystals (self-structured objects on large scales).

Boerlijst and Hogeweg's work is something altogether different. Unlike the work of Langton and others who interpret the spontaneous self-structuring in cellular automata as a form of evolution, Boerlijst and Hogeweg see this self-structuring as a "substrate for selection". Their work has specific importance in bringing hypercycles back into play as mechanisms that may facilitate prebiotic evolution; the self-generated and deterministic spatial structure in their models provides an environment within which interspecific group selection can operate in a robust way, ridding the system of parasitic molecules. More generally, the work is important in highlighting the need to examine evolutionary and ecological problems not in homogeneous environments, but in spatially structured environments that themselves can be generated by simple rules. As in most previous studies, Boerlijst and Hogeweg consider a situation where the self-generated spatial structure consists of spiral waves, but many other patterns, including deterministically generated chaos⁶, can arise. It is not clear to me whether Boerlijst and Hogeweg's models can ever generate chaotic substrates, nor what the consequences would be if they could, but my guess is that these kinds of questions are likely to be a growth industry. □

Robert M. May is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

1. Eigen, M. & Schuster, P. *The Hypercycle: A Principle of Natural Self-Organisation* (Springer, New York, 1979).
2. Maynard Smith, J. *Nature* **280**, 445-446 (1979).
3. Boerlijst, M. C. & Hogeweg, P. *Physica* **D48**, 17-28 (1991).
4. Nowak, M. & Schuster, P. *J. theor. Biol.* **137**, 375-395 (1989).
5. Gerhardt, M., Schuster, H. & Tyson, J. J. *Science* **247**, 1563-1565 (1990).
6. Hassel, M. P., Comins, H. N. & May, R. M. *Nature* **353**, 255-258 (1991).
7. Winfree, A. T. *When Time Breaks Down* (Princeton University Press, 1987).
8. Muller, S. C., Plesser, T. & Hess, B. *Physica* **D24**, 71-85 (1987).
9. Keener, J. P. & Tyson, J. J. *Physica* **D21**, 307-322 (1988).
10. Langton, C. (ed.) *Artificial Life: Studies in the Science of Complexity* (Addison-Wesley, New York, 1988).

Falsified teeth

Of all parts of the living human body, the teeth are most nearly dead. Their metabolic turnover is almost nil; so are their powers of healing or regeneration. Like dead wood, they submit passively to decay. They tolerate implants, like fillings and crowns, with no reaction at all. Their sole concession to awareness is the ability to feel pain.

More exasperating still, tooth-pain doesn't even warn of impending damage. It is futilely and pointlessly stimulated by the movement of fluid in the tiny tubules that run through the dentine. Heat, cold, drying or the osmotic pressure of sugar-based foods can all set up this flow and cause pain; decay does not. Tooth owners, says Daedalus, have every right to be aggrieved.

Structurally, teeth are inorganic apatite bonded and toughened by a cunning interweaving of organic protein. Tooth-decay attacks this protein. In this connection Daedalus recalls the traditional method of hardening biological specimens with formaldehyde solution. This cross-links their protein chains to give an inert, decay-proof, durable solid, chemically reminiscent of melamine resin. Similarly, he argues, treating a tooth with formaldehyde should bind its protein content into a tough, biologically inert and impermeable material — both stopping decay in its tracks, and sealing the pestilential tubules whose permeability permits the tooth to feel pain.

So Daedalus is inventing a cunning dental filling based on polyformaldehyde. This engineering plastic, better known as acetal resin, is kept from depolymerizing back to formaldehyde only by the most ingenious chemistry. A subtly inferior formulation should degrade very slowly, leaking formaldehyde monomer steadily into the dentine beneath. Formaldehyde is decidedly toxic, but can hardly be worse than mercury, a major component of many dental fillings. In any case, the tiny traces leaking from the filling will react almost instantly with the tooth protein. Soon the tooth will be fully crosslinked into a sealed, biologically inert structure, immune to pain and further decay. The polyformaldehyde filling can then be replaced by a more conventional one — perhaps of melamine resin — compatible with the new tooth material and bonding intimately to it.

Thus dentistry will be transformed. At the first hint of trouble, a tooth will not so much be filled with resistant material, as chemically converted into it. Thereafter, it will be totally inert. The long decline which now charts almost everyone's dental history will be arrested. In effect, we will end up wearing our own natural false teeth.

DAVID JONES

Maternal transfer of HCV

SIR — Hepatitis C virus (HCV) is the main causative agent for sporadic as well as parenteral cases of non-A, non-B hepatitis^{1,2}. It is not clear, however, whether the virus can be transmitted other than by blood transfusion. Mother-to-child transmission, which is the main route for the hepatitis B virus to establish new carrier states, has not been

However, five of the six nucleotides were on the third letter of the codons, resulting in only one amino-acid change. Such close similarity among HCV cDNAs in the family is explained if the grandmother transmitted the HCV to the mother at the mother's birth 27 years ago. This possibility is likely because the grandmother received a blood transfu-

These results indicate that the mother-to-child transmission of HCV does occur. The baby's hepatitis was self-limiting and transient. HCV infection in the mother and the grandmother was confirmed only by anti-p22 assay and by detailed analysis of the PCR products. It is possible that HCV transmitted from mother to child could result in transient or even subclinical hepatitis in babies which could persist for long periods of time. This may also explain the reason why a high prevalence of HCV infection among communities is maintained.

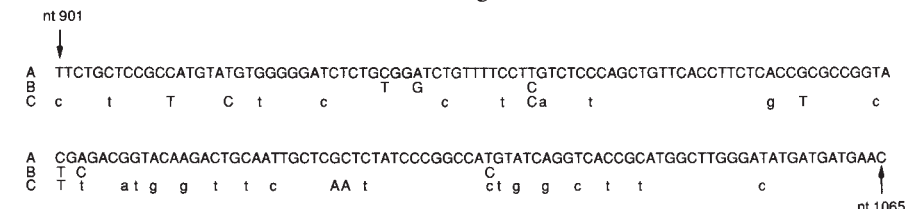
Y. INOUE
T. MIYAMURA

Department of Enteroviruses,
National Institute of Health,
Tokyo 141, Japan

T. UNAYAMA
K. TAKAHASHI

Department of Pediatrics,
Ashigarakami Hospital,
Kanagawa 258, Japan

I. SAITO
Laboratory of Molecular Genetics,
Institute of Medical Science,
Tokyo 108, Japan



Nucleotide sequences of a part of the HCV genome encoding the envelope protein. A, Complementary DNA derived from the baby and the mother; B, cDNA derived from the grandmother; C, cDNAs derived from four different HCV clones⁵⁻⁸. Only nucleotides different from A are shown in B and C. In C, capital letters indicate all four clones have the same nucleotides but different from A, lower-case letters show that at least one clone has indicated nucleotides different from A.

confirmed by existing anti-HCV assays. Among children under the age of 10, the prevalence of anti-C100, an assay detecting antibody to molecularly expressed HCV non-structural protein in yeast², is significantly low³. Such seroepidemiological findings even led to the notion that mother-to-child transmission of HCV is unlikely. If this is the case, it is unclear how HCV can persist in the community and maintain high carrier rates throughout the world (0.2–2.6%).

Here, we report a family in which HCV was transmitted from mother to child through two generations. A female baby developed acute, resolving type of hepatitis C and the diagnosis was confirmed by the anti-C100 assay. Although both her mother and grandmother were negative for anti-C100, they were considered to be healthy HCV carriers as they had the antibody against an HCV nucleocapsid antigen (p22) which was expressed in insect cells by a recombinant baculovirus⁴. Amplified HCV complementary DNA fragments were detected in the sera from the baby, the mother and the grandmother but shown by the polymerase chain reaction (PCR) method not to come from the father.

The nucleotide sequence of part of the HCV genome encoding the viral envelope protein was determined and analysed (see figure). The difference in the nucleotide sequences of this region among four different HCV cDNA clones reported in Japan is 16.3% (refs 5–8). We find, however, that the sequences of 165 nucleotides (901–1065) in the baby and the mother were completely identical, suggesting strongly that the HCV in the baby was derived directly from the mother. Six nucleotides (3.6%) were different from those of the grandmother.

sion 37 years ago and developed acute non-A, non-B hepatitis afterwards. Abnormal liver function was indicated when she delivered her first baby (the mother referred to above). In summary, as far as the region of the HCV genome we examined is concerned, the cDNA fragment from the baby is identical to that from the mother and more closely related to that from the grandmother than those from other Japanese isolates.

Neurovirulence factor

SIR — The ICP34.5 protein of herpes simplex virus type 1 (HSV-1) and its HSV-2 counterpart are potent neurovirulence factors^{1,2}: viruses defective in

HSV-1	193	VRFSPHVRVRLVWVWASARLARLARRGWARERADRFRFRVAAEAVIGPCLGPEARARALAR
HSV-2	166	VCFSPRVQVRHLVAVETAARLARLARRGWARERADRFRFRVAAEAVIGPCLGPEARARALAR
MyD116	549	VHFAEKVTVHFLAVWAGPAQAARRGPWEQFARDRSFARRIAQAEEKLGPYLTPDSRARAWAR

HSV Con	V-FSP-V-VRLV-W---AARLARLARRGWARERADRFRFRVAAEAVIGPCL-PEARARA-AR
Total Con	V-F---V-V-L-W---A-ARRG-W---DR-RF-RR-A-AE---GP-L-P---RARAR-AR

Comparison of amino-acid sequences of the 63-residue similar domains of HSV-1, HSV-2 and MyD116.

these genes require a 10⁵-fold increase in dose over wild-type to kill mice by intracerebral inoculation. We have found that the two protein sequences show most similarity (83% identical residues) in a 63-residue section near their carboxy termini. We have discovered that a mouse protein (MyD116; ref. 3) possesses a striking homologue of this 63-residue domain (see figure). Other parts of MyD116 are dissimilar to the HSV proteins, except for an acidic locus in each. No other database entries are similar to ICP34.5 or MyD116.

MyD116 is predicted from the sequence of a transcript expressed immediately during differentiation of a

myeloid leukaemic cell line induced by interleukin-6, and also in bone marrow⁴; it probably acts directly in controlling cell state, perhaps as a transcription factor or a modifier of cytoskeletal structure. ICP34.5 might act as a mimic or an antagonist of a MyD116-related cellular protein in infected brain cells. The re-

markable conservation of one region within otherwise distinct proteins from diverse sources suggests a functional module, other examples of which may emerge.

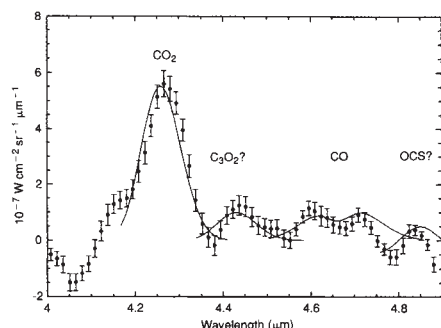
DUNCAN J. MCGEOCH
BARBARA C. BARNETT

MRC Virology Unit,
Institute of Virology,
University of Glasgow,
Glasgow G11 5JR, UK

1. Chou, J., Kern, E. R., Whitley, R. J. & Roizman, B. *Science* **250**, 1262–1265 (1990).
2. Taha, M. Y., Clements, G. B. & Brown, S. M. *J. gen. Virol.* **70**, 3073–3078 (1989).
3. Lord, K. A., Hoffman-Liebermann, B. & Liebermann, D. A. *Nucleic Acids Res.* **18**, 2823 (1990).
4. Lord, K. A., Hoffman-Liebermann, B. & Liebermann, D. A. *Oncogene* **5**, 387–396 (1990).

Carbon suboxide in comet Halley

SIR — Huntress *et al.* suggested¹ that carbon suboxide (C_3O_2) could exist in the nucleus of comet Halley, and that this species would be the progenitor of atomic carbon and carbon monoxide present in the cometary atmosphere. A C_3O_2 production rate of 0.03–0.04 times that of water would explain the observed



Reproduction of a part of the spectrum of comet Halley observed by the infrared spectrometer aboard the VEGA 1 probe⁴. Solid lines, synthetic band profiles for CO_2 , C_3O_2 , CO and OCS.

abundances and distributions of cometary C and CO. Huntress *et al.*¹ planned to look for C_3O_2 evidence in the Giotto mass-spectrometer data.

Another way to assess the presence of cometary C_3O_2 is to look for its infrared vibrational bands. Its ν_3 band at 4.43 μm is very strong² (band strength of 8,900 $cm^{-2} atm^{-1}$). The expected fluorescence rate of this band, at 1 AU from the Sun, is $1.0 \times 10^{-2} s^{-1}$, compared to 2.9×10^{-3} for the ν_3 band of CO_2 at 4.25 μm and to 2.6×10^{-4} for the ν_3 band of H_2O at 2.7 μm . If C_3O_2 had a mixing ratio of a few per cent, comparable to the mixing ratio of CO_2 , its ν_3 band would be stronger than the ν_3 band of CO_2 and very conspicuous in infrared cometary spectra.

We measured 2.5–5.0- μm spectra in comet Halley with the IKS experiment aboard the VEGA 1 probe^{3,4}. In addition to water, CO_2 , CO and H_2CO were detected with mixing ratios of 2.7%, 5% and 4%, respectively. OCS was tentatively detected with a mixing ratio of 0.7% or less. The figure shows that the ν_3 band of CO_2 is conspicuous, but that a weak feature of 4.45 μm may also be present. This must not be considered as a definite detection of C_3O_2 , as the feature may also be identified in other species (possibly $C \equiv N$ groups; ref. 4). If the feature is indeed due to C_3O_2 , its intensity corresponds to an abundance $[C_3O_2]/[H_2O] = 0.0011$ (see ref. 4 for method of calculation). This can be considered as an upper limit to the cometary abundance of C_3O_2 directly ejected from the nucleus.

Although the presence of a distributed source of C_3O_2 cannot be excluded, infrared observations show that parent C_3O_2 , if present, is less abundant than speculated by Huntress *et al.*

JACQUES CROVISIER
THÉRÈSE ENCRENAZ
MICHEL COMBES

Observatoire de Paris-Meudon,
F-92195 Meudon, France

- Huntress, W. T. Jr, Allen, M. & Delitsky, M. *Nature* **352**, 316–318 (1991).
- Pugh, L. A. & Rao, K. N. in *Molecular Spectroscopy: Modern Research* (ed. Rao, K. N.) Vol. II 165–227 (Academic, London, 1976).
- Combes, M. *et al. Nature* **321**, 266–268 (1986).
- Combes, M. *et al. Icarus* **76**, 404–436 (1988).

Tests for rodent polyphyly

SIR — Graur *et al.* suggested¹ that the order Rodentia may not be monophyletic, with guinea-pigs and their relatives being outside a clade consisting of Primates, Artiodactyla and the rodent suborder Sciuromorphi. They argue that the recognition of the guinea-pig and relatives as a separate order of mammals “may resolve or ameliorate many of the paradoxes associated with the evolution of guinea-pig genes”, such as parallel changes in guinea-pigs and birds at the molecular level and accelerated rates of gene evolution. Acceptance of this hypothesis requires the reinterpretation of many morphological characters as homoplastic between the two recognized suborders of Rodentia.

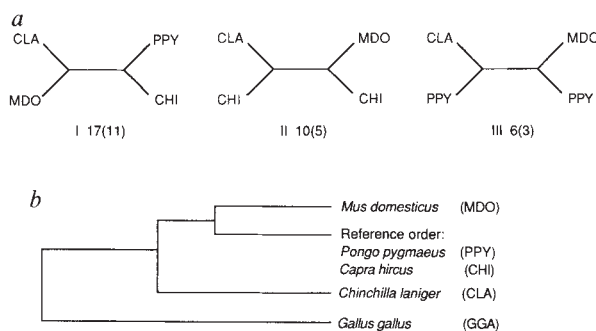
The order Rodentia represents a group of gnawing mammals that combine several morphological characters associated with specializations of the incisors, molar teeth and musculoskeletal features of the masticatory apparatus². Virtually all character systems examined have supported rodent monophyly, including those as diverse as the postcranial skeleton, placental and fetal membranes, carotid arterial patterns and

myology³. Although parallelism within the order may make determining relationships among families difficult, from a morphological perspective rodents constitute a distinct group of eutherian mammals and any taxonomic controversy has previously not included a question of their monophyly³. A hypothesis of rodent polyphyly therefore replaces the ‘molecular paradox’ resolved by Graur *et al.* with the problem of reinterpreting well-established morphological synapomorphies as homoplasies.

Rodent polyphyly, therefore, challenges what is known about the comparative morphology of this order. Resolution of this conflict between molecular and morphological data is critical for understanding the evolution and phylogenetic use of these different character systems. We have looked at nucleotide sequence data from the mitochondrial 12S ribosomal RNA gene for some of the taxa examined by Graur *et al.*¹. Only transversions were examined as these have been reported to change linearly with time throughout the eutherian radiation⁴ and for even older divergences⁵ and are least likely to be affected by multiple mutations. The most conservative of these transversions, those occurring in a ditypic state, were analysed separately.

We conducted two levels of phylogenetic analysis, the first to determine relationships among the eutherian orders and to illustrate the value of mitochondrial 12S rRNA gene sequences at the higher hierarchical levels, and the second to assign a root. In the first, both total and ditypic transversions support the arrangement grouping the two rodent suborders separate from the other represented mammalian orders (see figure). The support for this arrangement is significant at $P < 0.05$ for both total and ditypic transversions according to a binomial test⁶. Graur *et al.* also support this arrangement,

a, Three possible dichotomous arrangements of four mammals. The values indicate the number of informative transversions, and the number of ditypic transversions (in parentheses) supporting each arrangement. A ditypic transversion is defined as a sequence position with two species sharing the same purine, and the other two sharing the same pyrimidine⁷. Species abbreviations are as in b. b, Best tree found following Graur *et al.*'s four-taxon method with chicken as the designated outgroup. Ditypic transversions were used solely to compensate for the greater divergence between chicken and our mammalian ingroup. Due to difficulties in determining an optimal alignment, the most variable region of domain II, 5' of helix 14, was omitted in all cases. All parsimony analyses were conducted with the PAUP program (version 3.0K, D. L. Swofford). Published sequences available from GenBank.



suggesting that it can be rooted along the guinea-pig lineage supporting the polyphyly of Rodentia. Following their four-taxon method of analysis using two rodent suborders, a reference order and an outgroup, the same root as found by them is supported by our second analysis, with either reference order although by only one extra mutation apiece (*b* in the figure). The presence of statistical support for the eutherian orders implies that greater support for a root will be forthcoming when a more closely related outgroup is used. Comparable sequence data for a metatherian or edentate are therefore expected to provide a robust solution for the location of the root, and thereby a better test of rodent polyphyly.

MARC W. ALLARD
MICHAEL M. MIYAMOTO

Department of Zoology,
University of Florida,
Gainesville,
Florida 32611, USA

RODNEY L. HONEYCUTT
Department of Wildlife and Fisheries
Sciences,
Texas A&M University
College Station, Texas 77843, USA

1. Graur, D., Hide, W. A. & Li, W.-H. *Nature* **351**, 649–652 (1991).
2. Hartenberger, J.-L. in *Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis* (eds Luckett, W. P. & Hartenberger, J.-L.) 1–33 (Plenum, New York, 1985).
3. Luckett, W. D. & Hartenberger, J.-L. in *Evolutionary Relationships Among Rodents: A Multidisciplinary Analysis* (eds Luckett, W. P. & Hartenberger, J.-L.) 685–712 (Plenum, New York, 1985).
4. Miyamoto, M. M. & Boyle, S. M. in *The Hierarchy of Life. Molecules and Morphology in Phylogenetic Analysis* (eds Fernholm, B., Bremer, K. & Jönvall, H.) 437–450 (Elsevier, Amsterdam, 1989).
5. Mindell, D. P. & Honeycutt, R. L. *A. Rev. ecol. Syst.* **21**, 541–566 (1990).
6. Williams, S. A. & Goodman, M. *Molec. biol. Evol.* **6**, 325–330 (1989).
7. Lake, J. *Molec. biol. Evol.* **4**, 167–191 (1987).

Naming of the species

SIR — This is an appeal to molecular biologists presenting phylogenetic trees, in particular, that they give for each organism studied its genus and class (for example *Crepis*: Angiospermae), and not just its generic initial and specific name (as is done all too often nowadays). When we now read of *C. elegans*, *C. eugametos*, *D. discoideum*, *D. melanogaster*, *E. grandis*, *E. coli*, *H. haemophilus* and *H. sapiens*, not everyone can tell which are worms, flies, ferns, algae or bacteria. Details of species and strain should be presented in a table or in footnotes: not all conspecific individuals are genetically identical.

RALPH A. LEWIN

Scripps Institution of Oceanography,
University of California,
San Diego, La Jolla,
California 92093, USA

Lack's solution?

SIR — David Lack's paper¹ on the evolution of clutch size is one of the classics of evolutionary ecology. The Science Citation Index lists 398 references for the years 1985–89 to either this paper or Lack's book², which brought his ideas to a wider audience. Yet Lack's key insights on clutch size are indistinguishable from those of the unknown British ornithologist Eric B. Dunlop, published 34 years earlier³.

The nub of the hypothesis about optimal clutch size, which has become known as "Lack's solution"⁴, is that there is a trade-off between the number of eggs laid and the prospects of survival of individual chicks. If a bird lays too few eggs, it will leave few breeding offspring, yet if it lays too many, each young will receive too little food to have a good chance of survival. Thus there should be an intermediate clutch size which will maximize fitness. The greatest insight is that by laying fewer eggs a parent might leave more surviving offspring. Did Dunlop publish these insights? The answer is clear: he did. He was astonished at the frequency of starvation amongst late-hatching nestlings within broods of the rook, and this led him to make general comments about the many species of birds in which the starvation of late-hatching nestlings is common³.

... Now, if fewer eggs had been deposited and only the young hatched which would eventually have a reasonable chance of being reared, it would be greatly to the advantage of the individuals involved, for all the food brought by the parents would be divided amongst vigorous offspring and these would all receive a larger share than if there were more mouths to be fed, therefore these young would have a much better start in the world than those reared where there were young to be fed unprofitably and to some extent at the expense of the older nestlings. The adults also would benefit, for they would in all probability be relieved to a certain extent in their search for food, and yet would easily provide a larger amount for each individual of the smaller number of young.

This being so, it is clear that a reduction in the number of eggs laid would be of great value in such cases...

Dunlop then argues that in some species of Sulidae (gannets and boobies) a clutch of one has evolved from a clutch of two. Dunlop's prescience in the field of evolutionary ecology was thus as extraordinary as was the length of his sentences. But his life was short, as he was killed in the First World War at the age of 30, an acceptable excuse for not pursuing his ideas.

So should the idea of an optimal clutch size be called 'Dunlop's solu-

tion'? Perhaps, but such an historical retread may be unwelcome to the many evolutionary ecologists who have taken inspiration from Lack's work.

ROBERT D. MAGRATH
Division of Botany and Zoology,
Australian National University,
GPO Box 4, Canberra 2601, Australia

1. Lack, D. *Ibis* **89**, 302–352 (1947).
2. Lack, D. *The Natural Regulation of Animal Numbers* (Clarendon, Oxford, 1954).
3. Dunlop, E. B. *Brit. Birds* **7**, 105–114 (1913).
4. Godfray, H. C. J. *Oxf. Surv. evol. Biol.* **4**, 117–154 (1987).

Unpredictable earthquakes

SIR — In his commentary article¹, Geller stated that "the Japanese earthquake prediction plan (JEPP) has achieved no significant progress. Even now, no consensus exists on whether earthquake prediction is possible" and "Empirical earthquake prediction seems to have come to a dead end." The JEPP was established in 1965. Since then I believe that significant progress has been made in developing prediction technology and in earthquake research in general. The JEPP has been of great importance to related solid Earth sciences, particularly to seismology.

There have been many domestic and international meetings where significant progress in earthquake prediction research has been reported^{2–4}. Some earthquakes have actually been successfully predicted. Short-term prediction for the M7.0 Izu-Oshima earthquake in 1978 was issued by the Japan Meteorological Agency 1.5 hours before the event⁵.

There is a clear consensus that the state of the art in predicting earthquakes has progressed to the level that short-term prediction for a great interplate earthquake, the so called Tokai earthquake, can probably be made using the present instrumental networks, although neither the current state of knowledge nor the distribution of instrumentation permits useful prediction on a routine basis except for this earthquake. Precursors to an earthquake do not always occur but are very complex in their dynamics. Thus the practical approach to prediction is thought to be stochastic.

The JEPP emphasizes the observation of crustal activity over a long period of time. There have been many new discoveries of precursory phenomena of earthquakes. We found more than 10 different types of precursors to the 1978 Izu-Oshima event and to the 1989 seismo-volcanic activities, including an M5.5 earthquake and a volcanic eruption off the Izu peninsula. These included geochemical, geodetic, geoelectromagnetic and seismological anomalies. A proposed magma intrusion model quan-

titatively explained the anomalous crustal deformations associated with the 1989 activities. A plate model of the Kanto-Tokai region which is the scene of extensive seismic activity has also been constructed. These are only a few examples among many achievements. To accelerate the progress of earthquake prediction technology it is essential to increase the number of case studies using more intensified observation systems.

Geller's claim that there should be "external review" for the choice of research directions in the JEPP is reasonable. Not only do the formalities of an external review process need to be set up but there must also be a change in the general attitude of scientists. Careful consideration and discussion among scientists are important to solve these issues.

KAZUO HAMADA

National Research Institute for Earth
Science and Disaster Prevention,
Science and Technology Agency,
Tennodai 3-1, Tsukuba-Shi 305,
Japan

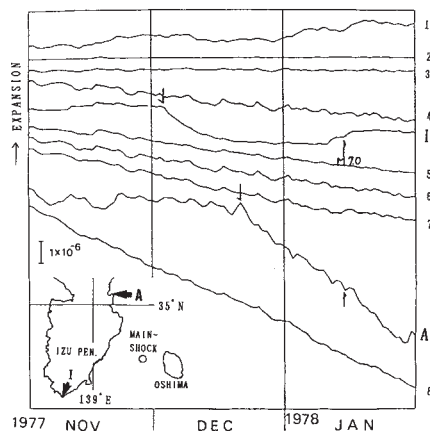
1. Geller, R. J. *Nature* **352**, 275–276 (1991).
2. *Proceedings of Earthquake Prediction Research Symposium* (Subcommittee of Earthquake Prediction, National Committee for Geophysics, Science Council of Japan and Seismological Society of Japan, 1980).
3. *Proceedings of Earthquake Prediction Research Symposium* (National Committee for Seismology, Science Council of Japan and Seismological Society of Japan, 1987).
4. *Report of the Coordinating Committee for Earthquake Prediction Vol. 1* (1969) – Vol. 45 (1991), Geographical Survey Institute, Ministry of Construction, Japan.
5. Japan Meteorological Agency in *Proceedings of 2nd Joint Panel Meeting of the U.J.N.R. Panel on Earthquake Prediction Technology 20* (Geographical Survey Institute, Ministry of Construction, Japan 1981).

GELLER REPLIES — Hamada's claim that a "short-term prediction for the M7.0 Izu–Oshima earthquake of 14 January 1978 was issued by the Japan Meteorological Agency (JMA) 1.5 hours before the event" is at variance with the facts.

The 1978 Izu–Oshima earthquake occurred in an area noted for earthquake swarms. Following 18 felt earthquakes ($M_{\max}$ 4.9) in a 3-hour period, the JMA issued the following statement¹ (my translation): "The present swarm events, which are somewhat larger than typical swarm events, are the largest since events in 1964 that caused a small amount of damage. Because there is a possibility that the present swarm earthquake(s) may cause some damage, you might consider taking precautions". The M7.0 earthquake occurred about 90 min after this statement. However, the JMA

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □



Volumetric strain records at 10 sites on the Pacific coast of Japan, near Tokyo. Ajiro (A) and Irozaki (I), shown on the inset map, are about 34 and 41 km from the epicentre, respectively. The other stations are at distances from about 60–200 km. The arrows in early December (Irozaki) and mid-December (Ajiro) indicate the beginning of alleged precursors. The arrows on 14 Jan indicate the time of the M7.0 earthquake. This figure (including the arrows exactly as drawn) is taken from the JMA data published shortly after the 1978 event.

statement, which merely cautioned that the earthquake swarm might continue, is not a 'short-term prediction', as it places bounds on neither the hypocentre, the time nor the magnitude². Indeed, neither the JMA nor other advocates of earthquake prediction in Japan have claimed the 1978 event as a successful prediction¹.

Hamada's claim of more than 10 different types of precursors for the 1978 Izu–Oshima event is apparently based on data summarized by Sekiya³. The volumetric strain recordings at Ajiro and Irozaki (see figure) are considered the best examples of precursors⁴, but no quantitative studies have shown these signals to be causally related to the M7.0 earthquake. The amplitude of the December–January signal (the alleged precursor) at Irozaki is an order of magnitude larger than the strain change for the M7.0 Izu–Oshima earthquake at Irozaki (which cannot be seen clearly in the figure, but can be seen clearly in data presented by Sacks *et al.*⁵). Because near-field strain dies off as r^{-3} , this suggests that the alleged precursor at Irozaki is the result of a source in the immediate vicinity of Irozaki, rather than a precursory source near the epicentre of the M7.0 earthquake, as otherwise a similar signal would have been observed on at least some of the other strain records in the figure. This argument also seems to apply to the Ajiro record.

The existence of many other precursors (water levels, radon and so on) for the M7.0 earthquake has also been

claimed. However, Shimazaki⁶ concluded that some of the apparent precursors are artefacts, and that reports of a similar pattern of temporal variation of different types of observations at different sites were groundless. Also, Turcotte⁷ has noted the absence of compelling physical arguments for viewing temporal changes of such phenomena as earthquake precursors.

Notwithstanding Hamada's comments, short-term prediction in Tokai is possible only if precursors occur, are observed, identified, and can be reliably analysed to constrain the parameters of the predicted earthquake. Turcotte noted that "large earthquakes . . . occur without systematic precursory phenomena"⁷. Hamada concedes that "precursors . . . do not always occur". No short-term seismic or strain precursors were observed before the M7.1 Loma Prieta, California, earthquake in 1989, despite extensive observational networks⁸.

Hamada claims that the present empirically oriented earthquake prediction programme has led to "significant progress . . . in earthquake research in general", but I think there would have been much more progress in the past 25 years if "earthquake prediction and basic research in seismology and other related fields [were] strongly linked"².

Hamada claims there is a "clear consensus that . . . short-term prediction for the Tokai earthquake can probably be made using the present instrumental networks". However, the spirited discussion and standing-room attendance at a recent special session "What do earthquake precursors signify?" at the Japan Earth and Planetary Science joint meeting in April demonstrated both a lack of consensus and an intense interest on the part of the scientific community.

Excepting my Commentary article, none of the references cited by Hamada appeared in a refereed journal. This epitomizes the closed nature of Japan's present earthquake-prediction programme and further underscores the need for an open review.

ROBERT J. GELLER

Department of Earth and Planetary
Physics,
Faculty of Science, Tokyo University,
Tokyo 113, Japan

1. Rikitake, T. *Earthquake Prediction and Earthquake Warning* (in Japanese) p211–212 (Centre for Academic Publications, Japan, 1979).
2. Geller, R. J. *Nature* **352**, 275–276 (1991).
3. Sekiya, H. *Coordinating Committee for Earthquake Prediction Report* **21**, 89–90 (1978) (in Japanese).
4. 20th Anniversary Report of Coordinating Committee for Earthquake Prediction p.12–13 (1990) (in Japanese).
5. Sacks, I. S., Suyehiro, S., Linde, A. T. & Snoke, J. A. *Tectonophysics* **81**, 311–318 (1982).
6. Shimazaki, K. *US Geol. Surv. Open File Rep.* **87**–591 p. 26–53 (1987).
7. Turcotte, D. L. A. *Rev. Earth planet. Sci.* **19**, 283–281 (1991).
8. *Lessons learned from the Loma Prieta, California, Earthquake of October 17, 1989*. US Geol. Surv. Circ. No. 1045 (1989).

Who's the noble beast?

Peter Singer

Factory Farming. By Andrew Johnson. Blackwell: 1991. Pp.272. £19.95, \$29.95.

WHEN humans first domesticated food animals, they took advantage of the fact that some animals offered a means of transforming foods that humans cannot eat into foods that humans can eat. Cattle grazed on grass, which our stomachs could not digest; but we could eat their flesh, or the milk they produced. Similarly, we took the eggs laid by hens which picked up insects and seeds that either did not appeal to our tastes, or that would not have been worth our while to collect.

The edge in efficiency that we gained by using animals in these ways remained the basis of animal food production for several millennia. Then, in the years after the Second World War, with no fanfare and no consultation with the public or its elected representatives, a new stage in our exploitation of other species began. With the advent of factory farming, animals were brought indoors and fed largely on grains or other products that were already suitable for human consumption.

This development was problematic in several ways. First, as Ruth Harrison argued in her pioneering *Animal Machines*, published in 1964 (Vincent Stuart), the degree of confinement frustrated the desires of the animals to behave naturally. Cruelty, she said, is acknowledged only where profitability ceases. In other words, as long as a hen keeps laying, a sow keeps breeding and a veal calf keeps putting on weight, it is of no consequence to the producer that the hen is unable to stretch its wings or scratch in the dirt, the sow cannot turn around or walk a single step, and the calf is kept deliberately anaemic so as to make its flesh the pale pink that gourmets prefer. Harrison's concerns about animal welfare raised the first serious doubts about the desirability of factory farming. A decade later, they were to play an important part in the development of the new animal liberation movement.

In 1971, a different attack on intensive farming was launched by Francis Moore Lappé in her bestselling *Diet for a Small Planet* (Friends of the Earth/Ballantine, revised edition, 1985). Lappé calculated,

for various forms of confined animals, the number of pounds of protein fed to livestock that it takes to produce one pound of protein for human consumption. The results were a shock to all who had believed that farming increases the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

intensive farming may not be the answer to mass production of food for humans. More than four pounds of protein are required for each pound of protein produced in the form of eggs; inefficiency at the expense of animal welfare, the environment and consumer taste.

availability of food for humans. The most efficient — or rather, least inefficient — forms of production turned out to be the egg and dairy industries, but even here it took more than four pounds of protein to produce one pound of protein in the form of eggs or milk. At the other extreme, more than twenty pounds of protein had to be fed to intensively reared cattle in order to produce a single pound of beef or veal. No wonder that Lappé called intensive farming "a protein factory in reverse". Her research had clearly demonstrated that animal-raising had entered a historic new phase, one in which it no longer produced any net gain in food available to human beings. For a small planet with a large population to feed, factory farming is, Lappé argued, a luxury too wasteful to be tolerated.

With the publication in 1980 of *Animal Factories* (Harmony Books, revised edition 1990), Jim Mason and I reiterated the problems of animal welfare and the wasteful use of land, but also pointed

out further problems for the environment, for the health of consumers and for the viability of rural communities. Factory-farmed animals cannot spread manure around the fields in a natural way; instead, the concentrated waste becomes a problem that often leads to the pollution of streams. There are also appalling odours for those living downwind of a large factory farm. To keep the crowded animals healthy and growing in their confined living space, a variety of drugs are routinely mixed into their feed, with unknown consequences

for consumers. Finally, because factory-farm systems are capital-intensive instead of labour-intensive, they contribute to the decline in rural employment.

Andrew Johnson's *Factory Farming* belongs in this critical tradition. Johnson approaches factory farming from a background in running a country-house hotel and restaurant on the Scottish island of Harris. Thus he has a strong interest in the quality and taste of the products of

animal farming, and is in no doubt that animals that range freely taste better when they arrive on the plates of patrons of his restaurant. Tasteless and flabby food, however, seems to be the least of the worries that face consumers of factory-farm produce. Johnson also points out that there is a risk of drug and hormone residues in intensively reared pork, beef and veal, while poultry may be covered in bacteria, and eggs carry the danger of salmonella.

Johnson also systematically investigates other concerns about factory farming. What he writes here is not always novel, but he brings earlier criticisms up to date. Thus, in addition to consumer concerns about food quality, drug residues and microbial contamination, he deals extensively with the impact of various systems of intensive farming on animal welfare and on the environment. The inclusion of fish-farming is particularly welcome. This relatively recent industry has been largely overlooked by other writers.

Holt Studios International/G. Roberts

Johnson sees the creation of an informed public as the best vehicle for change. As a contribution, he offers appendices that rate various kinds of animal products in terms of food quality, animal welfare and environmental impact. Genuinely wild game comes out best on each scale, whereas the main large-scale factory-farm products — battery eggs, intensively reared poultry and all nonfree-range pig products — are near the bottom on all three counts. Only farmed fish are, in Johnson's opinion, worse for the environment. In animal-welfare terms, intensively reared 'white' veal is still among the worst offenders, because much of it is imported from the continent, where veal calves are kept in dark, individual crates, even though such methods of rearing calves have been banned in Britain. Absolute zero on the welfare scale, however, is reserved for *pâté de foie gras*.

Factory farming has arrived with stealth, and grown to a size that makes it difficult to dislodge; but there are grounds for hoping that change is possible. Inroads have been made in several countries, largely on welfare grounds.

After banning individual crates for veal calves, the British government has now announced — in a development too recent for inclusion in *Factory Farming* — that it will also require individual sow stalls to be phased out. Sweden is now into the third year of a ten-year phase-out period that will eliminate all the most-confining systems of animal raising, and give every Swedish farm animal the right to an environment that allows freedom to move and meets minimal behavioural needs. In Switzerland, battery cages for hens will become illegal on 1 January 1992. Andrew Johnson's book shows that such moves are necessary not only because of our ethical obligations to animals, but also to preserve our environment and the quality of our food supply. If each nation would take what is most progressive from the others, we would already be a long way down the road towards ensuring that farm animals have some kind of life, and not only a mere existence. □

Peter Singer is director of the Centre for Human Bioethics, Monash University, Clayton, Victoria 3168, Australia.

Making the difference

Norman Myers

The Earth in Transition: Patterns and Processes of Biotic Impoverishment.

Edited by George M. Woodwell. Cambridge University Press: 1991. Pp.530. £30, \$49.50.

GEORGE Woodwell, president of the Woods Hole Research Center in Massachusetts, has been pondering the nature of biotic impoverishment since the mid-1950s when, using ionizing radiation, he destroyed all life in a sample plot of the Brookhaven Forest in New York State. Over the years (during one of which he served as president of the Ecological Society of America), he has investigated the consequences of many other types of human activities on natural systems, including forests, grasslands, tundras and aquatic habitats. He has concluded that we ought to focus not so much on the end results of biotic impoverishment — the graphic outcomes such as species extinctions — but on the nature of impoverishment itself, that is, the causes, courses and consequences of the Earth's biota in decline.

A worthy enquiry. What, for instance, is the character, extent and likely outcome of our current co-opting of 40 per cent of plant growth worldwide, as calculated by Peter Vitousek, Paul Ehrlich and their colleagues at Stanford Uni-

versity? What does our present understanding, meagre as it is, tell us about what could happen when there are — as there soon will be — twice as many people on Earth, demanding three times as much food and fibre? If we could co-opt 80 per cent or more of primary productivity, how would that affect the millions of other species that share the planetary ecosystem with us and how would that affect the human enterprise?

So too with the question of mass extinction of species. We hear much of how we are about to eliminate species in numbers that will match those of the 'great dying' at the close of the Cretaceous, when roughly half of all life forms are thought to have disappeared. How about mass elimination of genetic variability? During the next century many species that survive the biotic holocaust may well lose most of their subspecies, races and populations; genetically depauperate, they will then be all the less adaptable to drastically changed environments. Which phenomenon will prove to be the most detrimental to the Earth's biota and the many environmental services it supplies? Which will turn out to be the most disruptive for the functioning and ecological stability of natural communities? These questions carry momentous implications, with the effects of depletion probably persisting for at least a million years. Yet our understanding of the processes involved is rudimentary at best.

Hence the importance of this book, which presents papers from a late 1986 conference convened by Woodwell to

"define the major types of changes in the structure and function of natural communities exposed to chronic disturbance . . . Are [the changes] systematic, cumulative, predictable? At what point in their progress do they become identifiable? To what extent are they reversible?" The book does not seek to come up with diagnostic procedures for evaluating the state of our global experiment; rather it aims to "outline the broad patterns of the transitions underway." As such, it does a first-rate job.

The opening section surveys planet-wide patterns of impoverishment, with some emphasis on the effects of air pollution. Successive sections appraise the scope and scale of disturbances stemming from human activities. The reader is guided through a series of revealing case studies, examining forests in Australia, New Zealand, Pacific islands and Amazonia; xeric vegetation in Israel, the Great Basin of the western United States and California; Arctic ecosystems; marine communities in the Red Sea and the deep oceans; and freshwater systems in southern Florida and Ontario. The book concludes with a critical assessment of solutions, highlighting the need for shifts in macro-level policies for conservation of natural communities.

The list of contributors is impressive. It includes such prominent scientists as Herbert Bormann, Margaret Davis, Dieter Mueller-Dombois, Robert Repetto, Zev Naveh, Walter Westman, Fred Grassle, John Cairns, Gus Speth and Donella Meadows. Most of them present an excellent account of their research area, and each paper is introduced with an incisive scene-setting assessment by Woodwell. But one wonders which of these luminaries was so tardy in submitting his or her revised paper that it has taken four years for the book to see the light of day. Do they not recognize the ultra-urgency dictated by the speed of change overtaking the world outside their laboratory windows?

All in all, this is an exceptionally competent book, the only one of its sort to achieve a systematized, also systemic, insight into the question of how far and how fast we are reducing the biotic diversity of the Earth. Although no more than exploratory in certain respects, and offering few conclusive findings, it is a splendid pioneering effort to address one of the great environmental issues of our time. It eschews the (unduly?) dramatic assertions of apocalyptic analysts who proclaim the 'death' of ecosystems. Rather, it concentrates on those factors — less spectacular but more illuminating — that bespeak the progressive ill-health of our biota. □

Norman Myers, a consultant in environment and development, is at Upper Meadow, Old Road, Headington, Oxford OX3 8SZ, UK.

Comparatively speaking

D. J. Siebert and C. J. Humphries

Phylogeny, Ecology and Behavior: A Research Program in Comparative Biology.

By Daniel R. Brooks and Deborah A. McLennan. University of Chicago Press: 1991. Pp.434. \$45, £35.95 (hbk); \$21, £16.75 (pbk).

The Comparative Method in Evolutionary Biology. By Paul Harvey and Mark Pagel. Oxford University Press: 1991. Pp.239. £30 (hbk), £12.50 (pbk).

TAXONOMY — the science of the classification of organisms — has come a long way since about 1960. Branching diagrams of possible relationships among organisms are now commonplace, giving biologists new opportunities to investigate evolutionary problems in terms of closeness of descent (phylogenetics). These volumes aim to show how disciplines outside taxonomy might take advantage of phylogenetics. Brooks and McLennan outline a broad research programme for the study of the ecology of organisms through history, making explicit use of phylogenetic trees. Harvey and Pagel complement this approach with a discourse on statistical treatment of adaptation with regard to phylogenetic relationships among taxa.

Brooks and McLennan's recipes are demanding and problematic, but might be worthwhile pursuing. Fundamentally, cladistic analyses of species are required whether they are used for a study of speciation, adaptation or ecological association. But the number of groups of organisms for which trees are available is small, and the number of trees available for different groups within a given ecological community is even smaller. Brooks and McLennan rule out the use of written classifications as the source of phylogenetic trees. This is harsh because some classifications reflect phylogenetic relationships exactly.

For studies of speciation, Brooks and McLennan demand the complete enumeration and inclusion of all species (including those extinct) of a clade (or monophyletic group). A clade consists of individuals descended from a common ancestor. But this prescription will not work, for the simple reason that complete enumeration of species in a clade is never knowingly achieved. Furthermore, the authors use the hennigian model of speciation by which a species gives rise to two daughter species, the ancestor becoming extinct in the process. But it is impossible to demonstrate the existence of ancestors, so how are extinct taxa, required by the hennigian model, in-

cluded in the analysis? The model was used axiomatically by William Hennig in the development of phylogenetic techniques, but was shown later to be irrelevant to cladistics. So what really is its use, especially as it cannot be established empirically?

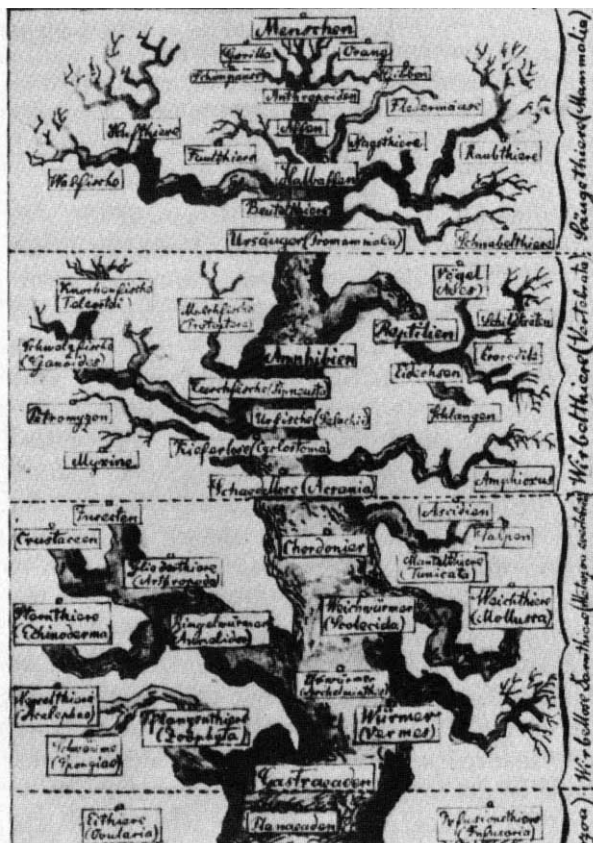
The authors analyse different groups

The main advance has been the development of methods based on explicit evolutionary and statistical models." But the real contribution is surely support for the use of phylogenetic systematics in population biology, which amounts to identifying the level in the phylogenetic hierarchy to which features of organisms

properly apply, and to elucidating the correct number of degrees of freedom in such comparative studies. Consideration of phylogenetic history certainly sharpens population biology and enriches the subject, as Harvey and Pagel amply demonstrate.

Phylogenetic history ensures that some species are not independent for statistical purposes, an ingredient necessary in Harvey and Pagel's recipes.

Model-independent pattern analysis, the most general phylogenetic taxonomic method, provides cladograms as the framework for identification of statistically independent taxa. Ironically, Harvey and Pagel completely dismiss 'pattern' cladistics because this method is not dependent on models. The authors require (admittedly inadequate) models because such models allow statistical treatment. Their defence of such models is weak,



Branching out — today's phylogenetic trees are far removed from this portrayal of the 'tree of life' by Ernst Haeckel, Darwin's German translator. (Reproduced from *The Encyclopedia of Evolution*, Facts on File, New York.)

of organisms associated with each other (parasites on hosts, communities of fishes in North American streams and so on) by comparison of their respective cladograms. These historical associations are approached through a method known as Brooks' parsimony analysis. The authors present this analysis as being superior to all other approaches by ignoring criticism and representing other methods as flawed. But Brooks' parsimony analysis is itself flawed because it fails to recognize the problems of polyphyly and paraphyly in many different associations and widespread taxa. Also, it is model-dependent and less general than its main competitor, component analysis.

Harvey and Pagel outline recipes of a much less general nature. Their aim is to cloak neodarwinian schemes with the rigours of phylogenetics and statistical analyses. They claim that "the comparative method has, however, changed radically in recent years, and this book is about a new kind of comparative study.

being no stronger than the assertion that methodologies not using them are likely to be based on even more unrealistic implicit models. Maximum-likelihood techniques for phylogeny reconstruction may have statistical properties that are considered to be advantageous. But when the underlying models are unrealistic, the advantageous properties will ensure compliance only with the initial inadequate model, rather than with the most defensible hypothesis.

Harvey and Pagel's general discussion on taxonomy is naive, and their concept of rigorous phylogenetics amounts to hollow rhetoric: "Our conclusion is that cladistically defined taxonomies are usually the most suitable, but that some of the procedures commonly used by cladists can be improved upon to provide better estimates of *true phylogenies*" (our italics). Even perusal of their taxonomic references is revealing for inappropriate citations, absent critical papers and substitution of primary sources with weaker secondary literature.

Perhaps this is understandable, because neither Harvey nor Pagel is a taxonomist. The issue of 'true phylogenies' is a dead topic, and their claim that molecular taxonomies are better estimates of phylogeny was laid to rest a while ago.

Comparative biologists of all persuasions might enjoy reading these books and they might benefit by seeing how to incorporate phylogeny into their work. Taxonomists might glean ideas on how to put the results of their labours to use outside taxonomy. The broader application of phylogenetics is yet to be fully explored, and improvements in methodologies, even new ones, are to be expected. The protocols put forward in both books should be looked upon as a beginning, but not necessarily copied.

The truth remains that neither of the books is really about taxonomy. Rather,

each is propaganda for the uses to which taxonomy might be put.

Neither book presents a full or critical discussion of phylogenetics, and a reader unfamiliar with the field should not expect to gain a full understanding from them. Taxonomy is a burgeoning science and there is no complete guide to it. Primary sources must be consulted for a better understanding of the techniques described. Of the two, Brooks and McLennan's books, should be read first. It is wider in scope, richer in possibilities, better referenced and presents a stronger case for phylogenetics. Those interested in studying and testing for adaptation should consult both. □

D. J. Siebert (Department of Zoology) and C. J. Humphries (Department of Botany) are at the British Museum (Natural History), Cromwell Road, London SW7 5BD, UK.

Watch this space

Rainer Weiss

The Detection of Gravitational Waves. Edited by David G. Blair. Cambridge University Press: 1991. Pp. 481. £95, \$125.

THE prospect of detecting gravitational radiation from astrophysical sources is an exciting one. It promises new insights into the physics of highly relativistic (non-newtonian) gravitation at the sources and will provide a way of observing phenomena in the Universe that cannot be examined by techniques based on the more established electromagnetic astrophysics. Gravitational waves, of sufficient strength to be detected, will arise from the coherent motions of massive objects moving at relativistic speeds. The waves, once emitted, weakly couple to intervening matter and travel through the Universe unimpeded, carrying with them the imprints of their sources. One could expect to 'see' the actual dynamics in a stellar collapse, or the oscillations in the nuclear matter after the collision of a pair of neutron stars, or the ringing of the highly curved geometry around a black hole following a perturbation, or even to be able to probe the earliest moments of the cosmic explosion. All of these are invisible to electromagnetic astrophysics.

But the ability of gravitational waves to move through the Universe unscathed also means that the radiation is difficult to detect. The main part of this book is dedicated to a survey of the techniques and strategies that have been developed (up to late 1989) to detect the minute distortions (strains) of space in a gravitational wave. Because the electromag-

netic sky provides few clues to the strongest sources of gravitational waves, the range of plausible wave strengths and the event rates at which they occur are enormous. But by the same token this uncertainty affords an unrivalled opportunity to learn something really new. The strongest sources could produce impulsive strains as large as 10^{-19} , although such sources are probably rare. With a strain sensitivity of 10^{-21} , searches are possible for a variety of sources encompassing enough other galaxies to make it likely that as many as a few sources per year will be found. At a sensitivity of 10^{-22} , especially at observation frequencies of 100 hertz and lower, it is now certain that the orbital decay and coalescence of binary neutron stars will be observed at the rate of a few per year by sampling as far as 0.5-giga light years. Although this applies to impulsive sources, the most likely to be measured first, periodic sources such as pulsars in our own galaxy and a stochastic background of gravitational waves are other candidates for observation.

The notion of measuring a strain as small as 10^{-19} (considered to be a large amplitude astrophysically) in times as short as a hundredth to a thousandth of a second is daunting to most scientists. It means measuring a distance of 10^{-17} centimetres in a metre, a small fraction of a nuclear diameter, or 4×10^{-14} centimetres in 4 kilometres (just about a nuclear diameter), and this in a macroscopic system. The current state of the art, developed in part by many of the contributing authors in this volume, is a strain sensitivity of about 10^{-18} . This is the result of dogged, cunning and inventive work, much of it described in the book, by a group of dedicated people. They have concentrated, as few other scientists have had to, on the sources of noise and ultimate metrological limits of

their apparatus. Their persistence is all the more remarkable because in this field there has been a lack of interim results that encourage and characterize technical development elsewhere. For here the gap between apparatus performance and required sensitivity has been too large.

Technical development has now brought the field to the threshold of 'doing science', but to close the gap will require substantial financial investment. The most promising detection systems are the interferometric detectors which need to be scaled from tens of metres to kilometres to gain the necessary sensitivity. The scaling does not come free; it requires the construction of large baseline interferometers incorporating, in particular, large vacuum systems, which transform the 'table top' style and cost of this research to a level corresponding to a medium-size space mission or nuclear-research facility. Not big science, such as an elementary-particle accelerator, but nevertheless sufficiently large so that research groups once fiercely independent, are having to band together in both Europe and the United States to make an organized effort.

With these increased stakes, it is all the more important that researchers in other areas, not merely the practitioners, should come to understand the science and the technology of the field so as to make their own judgements of its merits. This understanding will also equip them to enjoy the development and discoveries of the field. Unfortunately, this book does not serve this critical function. It provides an informative and useful presentation of many of the technical issues associated with the field, written for experts and in the language of the field. But it does not provide the overview or the rationale for the added investment being asked from society in a way that could be understood by all scientists. In short, a missed opportunity. □

Rainer Weiss is at the Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

New in paperback

■ In *Global Ecology*, Colin Tudge provides an informal introduction to the interactions of living organisms and their environment. Amply illustrated in colour, the volume covers all areas of ecology. Published in Britain by The Natural History Museum, price is £6.95 and in the United States by Oxford University Press, \$29.95 (hbk only).

■ A second edition of *Tropical Rain Forest Ecology* by D. J. Mabberley has been published by Blackie. The book concentrates on the trees and other plants that comprise the rain forest, but this edition has allowed the inclusion of more zoological literature as well as the update and expansion of other text. Price is £17.95.

Airborne observations of the physical and chemical characteristics of the Kuwait oil smoke plume

D. W. Johnson*, C. G. Kilsby*[†], D. S. McKenna*, R. W. Saunders*,
G. J. Jenkins*, F. B. Smith[‡] & J. S. Foot[§]

* Meteorological Office, Meteorological Research Flight, Royal Aerospace Establishment, Farnborough, Hampshire GU14 6TD, UK

[†] Present address: Civil Engineering Department, University of Newcastle, UK

[‡] Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ, UK

[§] Meteorological Office, Remote Sensing Instrumentation, Royal Aerospace Establishment, Farnborough, Hampshire GU14 6TD, UK

Airborne measurements in the densest part of the smoke plume at about 120 km from the burning wells in Kuwait in late March 1991 showed typical particulate mass densities of 500–1,000 $\mu\text{g m}^{-3}$, mixing ratios of 500–1,000 p.p.b.v. of sulphur dioxide and 30–60 p.p.b.v. of nitrogen oxides. One thousand kilometres from Kuwait, ozone concentrations in the plume exceeded background levels by about 50 p.p.b.v. The oil burn rate was estimated from sulphur fluxes to be 3.9 ± 1.6 million barrels per day. Significant amounts of smoke were observed only below 5,000 m altitude, and the measured attenuation of solar radiation by the smoke was similar to those assumed in recent assessments.

ABOUT 600 naturally pressurized oil wells were set alight in Kuwait in late February 1991, injecting massive quantities of smoke, sulphur dioxide (SO_2), unburnt hydrocarbons and nitrogen oxides ($\text{NO}_x = \text{NO} + \text{NO}_2$) into the atmosphere. Assessments^{1–3} of the impact of these pollutants are subject to large uncertainties in the magnitude and the characteristics of the emissions and the resulting plume. To investigate the chemical and particulate composition of the smoke, and its effect on visible and thermal radiation, the C-130 aircraft of the Meteorological Office Research Flight made 57 h of observations in eight flights into and around the plume in March 1991. Here we report

the first airborne measurements of the plume both close to the source (up to 200 km) and in the far field (1,000 km from source).

Meteorology and plume behaviour

Typically, the smoke from individual oil-well fires combined and rose so that the plume top reached ~4,500 m, often with shallow convective clouds forming on the windward edge of the plume. There was a marked vertical wind shear leading to differential advection of the plume, with, for example, the lower part of the plume being transported southeastwards down the Gulf, with the upper part moving over Iran.

There has been speculation that, because of radiative self-heating, aerosol may be lofted⁴ into the stratosphere, where its long lifetime would allow transport over considerable distances. Visual observations showed the plume top to be well defined with a maximum height of 5,000 m. On one or two occasions a detached smoke layer was seen a few tens of metres above the main top. These observations in springtime show that the smoke was confined to the lower half of the troposphere and significant amounts were not being lofted into the stratosphere, in agreement with recent assessments^{2,3}.

Particulates and chemistry in the near-field

Vertical profiles and cross-wind horizontal runs through the plume were carried out on 28 March 1991, 120 km downwind of the source (defined as 29° N, 48° E) south of Kuwait city. This distance was chosen as it was close enough to the source areas that all emissions could be measured while not so close that the inhomogeneity of the individual sources rendered the measurements unrepresentative.

A passive-cavity aerosol spectrometer probe (PCASP) was used to measure smoke particle concentrations and deduce the

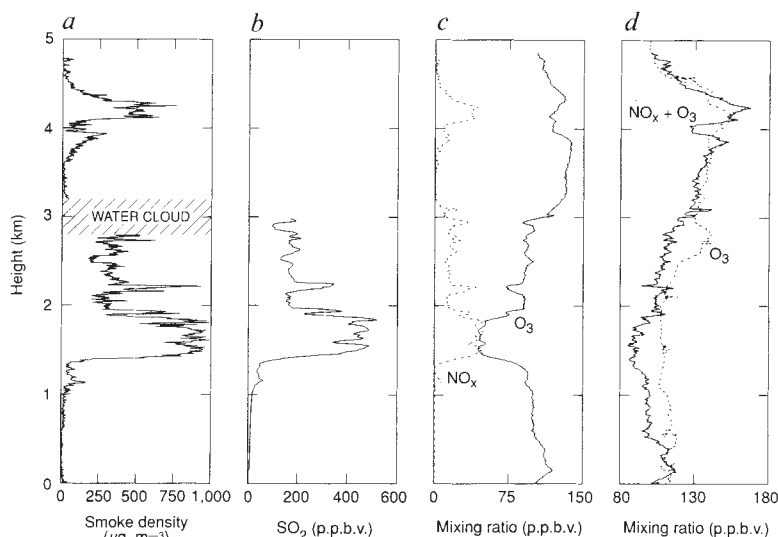


FIG. 1 Vertical cross-section of the plume made in a descending profile on 28 March 1991, 120 km from the source. *a*, Smoke density ($\mu\text{g m}^{-3}$); *b*, SO_2 (p.p.b.v.); *c*, O_3 (solid line) and NO_x (dashed line); *d*, O_3 outside the plume (solid line) and $\text{O}_3 + \text{NO}_x$ in the plume (dotted line).

TABLE 1 Hydrocarbon analyses of plume samples

Species	Sample A 30 March 191		Sample B 30 March 91		Sample C 29 March 91	
	Mixing ratio (p.p.b.v.)	Ethane ratio	Mixing ratio (p.p.b.v.)	Ethane ratio	Mixing ratio (p.p.b.v.)	Ethane ratio
ethane	25.111	1.000	5.440	1.000	1.330	1.000
ethene	13.030	0.519	1.851	0.340	0.034	0.026
propane	19.952	0.794	2.345	0.431	0.453	0.340
propene	1.637	0.065	0.220	0.040	0.026	0.020
benzene	1.057	0.042	0.215	0.040	0.044	0.033
toluene	1.309	0.052	0.203	0.037	0.025	0.019
<i>n</i> -hexane	5.052	0.201	0.409	0.075	0.013	0.010
NMHC	339.6	13.53	45.92	8.44	6.266	4.71

Analyses of 1.6-litre stainless steel flasks, by gas chromatography-flame ionization detection. Sample A: most polluted sample, 120 km from source; sample B: typical polluted sample at a similar distance; sample C: typical far-field sample (1,000 km from source).

mass of smoke particles per unit volume (smoke density). Continuous measurements were made of SO_2 , NO_x and ozone (O_3) with accuracies of $\pm 5\%$, $\pm 10\%$ and $\pm 5\%$ respectively. Stainless steel flasks were pressurized with ambient air and subsequently analysed for C_2 – C_8 hydrocarbons.

Figure 1 shows pollutant concentrations, together with the sum of O_3 and NO_x mixing ratios, from a descending profile carried out alongwind through the centre of the plume. Also shown is an O_3 profile obtained earlier outside the plume. The plume had a complex vertical structure, made up of two main layers: an upper one with a top at 4,600 m and base at 3,600 m, and a deeper, denser lower layer with a top at 3,200 m and base 1,000 m. The lower layer was capped by thin (200 m) altocumulus with droplets of high concentration (270 cm^{-3}) and small effective radius ($2.5 \mu\text{m}$). The upper layer was advected eastwards while the lower layer was advected southeastwards down the Gulf, trapped by a weak temperature inversion at the top of the liquid water cloud.

Seven horizontal, cross-wind runs over the same ground position were carried out at 1,420 m, 1,520 m (twice), 1,980 m, 2,440 m and 4,420 m (twice). Figure 2a–d shows measurements from a run through a 30 km-wide plume (near the centre of the

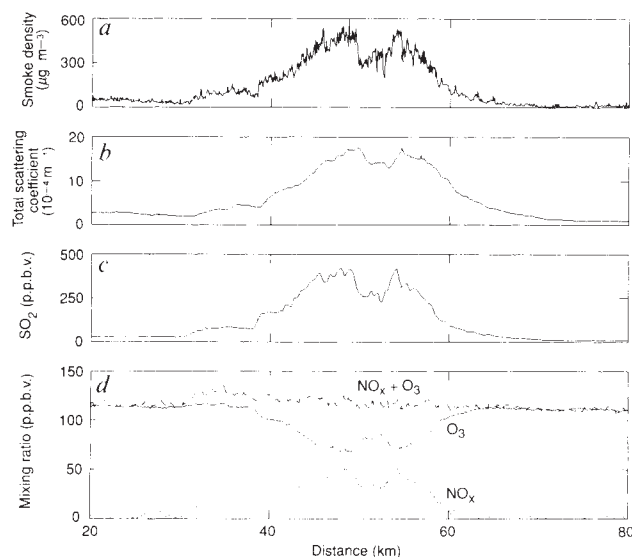


FIG. 2 Horizontal cross-section through the plume at 2,440 m on 28 March 1991. a, Smoke density; b, total scattering coefficient at $0.48 \mu\text{m}$ as measured by an integrating nephelometer; c, SO_2 ; d, O_3 (solid line), NO_x (short-dashed line) and $\text{O}_3 + \text{NO}_x$ (long-dashed line).

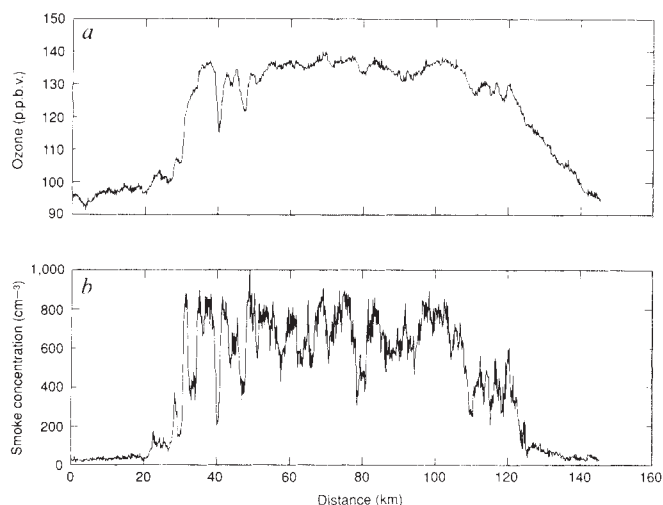


FIG. 3 Horizontal run at an altitude of 4,400 m through the area of smoke ($\sim 1,000 \text{ km}$ from Kuwait) on 29 March 1991 showing a, O_3 mixing ratio; b, smoke concentration.

lower layer at 2,440 m) of pollutants, total scattering coefficient and $\text{NO}_x + \text{O}_3$. In both the vertical profiles and horizontal runs all the plume indicators (SO_2 , NO_x , particulates and scattering coefficient) were highly correlated; this correlation was observed over several days of measurements close to the source. In the lower layer, NO_x and O_3 mixing ratios were negatively correlated, indicating gas-phase titration of O_3 by NO , but their sum showed little horizontal variation (Fig. 2d), indicating little loss of O_3 either by reactions with smoke particles or in the oxidation of SO_2 to sulphate. In the upper layer, the observed correlation between NO_x and O_3 was zero or positive, indicating that photochemical oxidant production had already been initiated.

Analyses of hydrocarbon samples taken in different parts of the plume are summarized in Table 1. The flask that showed the highest mixing ratios of hydrocarbons (sample A) was filled at an altitude of 2,600 m in the densest part of a smoke plume on 30 March 1991, 120 km from source. The aggregate of the carbon contained in all the identifiable non-methane hydrocarbons (NMHC) implies the presence of at least 340 p.p.b.v. of carbon as unburnt hydrocarbons. Assuming unidentified species having a flame ionization response are hydrocarbons, we infer a carbon content of over 460 p.p.b.v. A second flask filled on the same day (sample B) at an altitude of 1,500 m near the bottom of the plume was a more typical sample; compared with a UK urban air sample, the maximum mixing ratios of benzene and toluene in the plume sample were lower, but that of *n*-hexane was substantially greater.

Two supersaturation cloud condensation nuclei (CCN) spectra taken on 28 March 1991, one just below the top of the upper smoke layer at 4,500 m and the other in clear air at 5,200 m, indicate the plume contains an order of magnitude more CCN for a given supersaturation. For example, at 0.6% supersaturation, CCN concentration increased from 140 cm^{-3} to $2,000 \text{ cm}^{-3}$. In the thicker parts of the plume, CCN were so numerous ($> 4,000 \text{ cm}^{-3}$) that the counter became overloaded.

Far-field measurements

On 29 March 1991, satellite images showed smoke over central Saudi Arabia, some distance from Kuwait, oriented southwest-northeast and being advected to the northeast. The O_3 mixing ratios and particle concentrations observed in a horizontal run through this area of smoke are shown in Fig. 3a and b. In this more distant plume, it is likely that the ozone, enhanced by 40–60 p.p.b.v. above the already high values in the adjacent clear air, is formed from ozone precursors (NO_x and NMHC) released

from the oil fires. Sample C (Table 1), taken in this distant plume, has a different distribution of hydrocarbons from the hydrocarbon samples collected near the fires. It still has relatively high mixing ratios of the longer-lived hydrocarbons, such as ethane and propane, but short-lived hydrocarbons such as ethene and propene are substantially reduced. The low mixing ratios of shorter-lived species suggest that considerable OH degradation has occurred, as is consistent with the enhanced ozone mixing ratio. Although the NO_x mixing ratios in the plume were below the limit of detection (5 p.p.b.v.), further ozone production remains possible because mixing ratios of only several parts per 10^{12} (p.p.t.v.) are required to generate ozone⁶. The absence of many of the more reactive and therefore short-lived hydrocarbons means, however, that further ozone production will proceed more slowly.

The observations of this mature area of smoke suggest an interpretation of some of the clear-air observations in the near field. During many of the vertical profiles made near the source region but outside the visible plume, ozone mixing ratios much greater than the normal free tropospheric background (30–40 p.p.b.v.) were encountered (for example, Fig. 1*d*), associated with slightly increased particle density. The enhanced ozone mixing ratios observed in the clear air were probably formed from NO_x and NMHC in smoke emitted several days earlier that remained near, or returned to, the source area.

Morphology of smoke particles

During flights in the smoke plume, ambient air was drawn through polycarbonate filters which were subsequently studied with a scanning electron microscope (SEM). Figure 4*a* shows a typical smoke particle from a filter exposed 180 km from the source, on 30 March 1991. The smoke particles are composed of spherules of diameter $\sim 0.1 \mu\text{m}$, which form chains of hundreds in number and several micrometres in length, as observed in other studies^{4,7}. Particles on filters exposed more than 1,000 km from Kuwait (Fig. 4*b*) have a different appearance from those from the near field; although the spherule size is similar in both cases, the spherules in the far field are more

TABLE 2 Estimates of strength of oil-well sources

Component	This study (Mt yr ⁻¹)	Ref. 2 (Mt yr ⁻¹)	Ref. 1 (Mt yr ⁻¹)	Ref. 3 (Mt yr ⁻¹)
Total burn rate	202.5	80	63	161
Carbon as gas	161.2*	60	63	—
Fine carbon particulate	6.4†‡	5	5.8	16.1
Sulphur	6.1§	2	—	—
Nitrogen in oxides	0.42†	0.5	—	—

Estimates derived from observations and, for comparison, those assumed by recent assessments^{1–3}. The mass of hydrogen contained in H_2S and hydrocarbons in the fuel has been taken into account when determining the different components of the emissions.

* Deposits from large partially burned oil droplets near the well heads may reduce this total.

† Assuming the same emission factors for the southern and northern fires.

‡ This quantity could be up to 2–3 times higher with a commensurate reduction in carbon as gas.

§ Assuming a 1.5% sulphur content for the northern fields which contribute 15% overall to the emissions.

likely to be closely packed into near-spherical clusters. Similar behaviour has been noted⁸ in laboratory experiments, in air aged at high humidity.

The PCASP determines the number and size distribution of particles by passing them through a laser beam and measuring the scattered light intensity, using a relationship between intensity and size derived from experiments with latex spheres of known radius. The smoke density is calculated by integrating the size distribution assuming (1) that collection efficiency is 100%, (2) that the nonspherical particles in the plume scatter the same amount of light as spheres of the same mass (we conclude from previous experimental⁷ and theoretical⁹ work that, for the particle shapes encountered, this assumption is reasonable) and (3) that the density of the smoke particles is 1 g cm^{-3} , in line with Sokolik¹⁰. Support for this approach comes from the mass loading of the filter samples, which indicate that the absolute determination of smoke density from the PCASP is not grossly in error, and from the linear relationship between

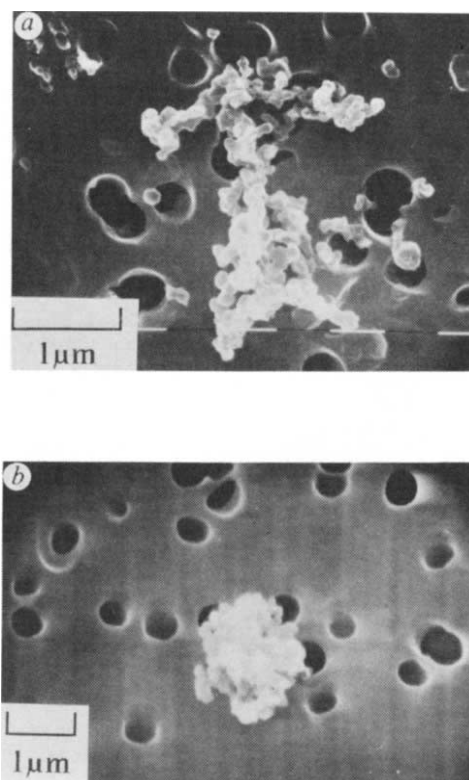


FIG. 4 Scanning electron micrographs of smoke particles collected on filters exposed in the plume on *a*, 30 March 1991, 180 km from the source, and *b*, 29 March 1991, $\sim 1,000$ km from the source. The filters were mounted on an extendible boom 0.4 m from the aircraft skin and the air sampled isokinetically¹⁴. The dark holes in the photograph are the pores in the filter.

SO₂ concentrations and smoke density (see below) which confirms that the instrumental response to smoke density is linear. Nevertheless, the uncertainties in these assumptions lead us to believe that the absolute values of smoke density could be up to 2–3 times as high as those shown. The relative variations in the plots are still valid.

Emission estimates

The amount of sulphur emitted was estimated from the flux of sulphur through a crosswind plane, determined from the product of sulphur concentration and the wind component perpendicular to the plane. Sulphur dioxide measurements from the horizontal runs and vertical profiles through the plume were used, together with the winds measured by the aircraft. Above 3,000 m, where the SO₂ monitor will not operate, PCASP measurements were used to derive proxy SO₂ values through the linear relationship evident in Figs. 1 and 2.

On 28 March 1991, when virtually all the plume ~120 km from the southern oilfields was sampled, the estimated sulphur emissions were 5.7 Mt yr⁻¹. Assuming a value of 3.3% for the sulphur content of the southern oilfields (R. D. Small, personal communication), this source strength corresponds to an oil burning rate of 172.1 Mt yr⁻¹. We assume, based on previous estimates^{11,12}, that the southern oilfields account for 85% of the total burning rate; this implies a total rate of 202.5 Mt yr⁻¹ or 3.9 × 10⁶ barrels per day. An estimate for 30 March 1991 gives a similar burning rate. Our subjective estimate of the error in this source strength, based on the temporal evolution and spatial variability over the 3 h measurement period, is ±40%.

Using the relationships between concentrations of SO₂, NO_x and particulates, we estimate the source strength of NO_x and particulates from the sulphur source strength (Table 2). Carbon

as gas (the mass of carbon contained in CO, CO₂, CH₄ and NMHC) is estimated from the residual mass after the mass emissions of sulphur, nitrogen and particulate carbon are subtracted from the total mass emitted and an allowance is made for the hydrogen content.

Effect of the smoke on radiation

Figure 5 shows the effect of the smoke on the radiation during a horizontal run on 31 March 1991 through the centre of the plume, 700 m below its top, ~65 km from source. The total shortwave (0.3–3 µm) downwelling flux was reduced from its clear-sky value of 800 W m⁻² outside the plume to zero in the centre, and corresponding changes were seen in the infrared and microwave regions.

The single scattering albedo, ω_0 , of the smoke was estimated in two independent ways. First, we measured hemispheric reflectances of the plume top, using upward- and downward-looking pyranometers, to be 5–8% over a range of solar zenith angles (25°–48°). The Ginzburg and Sokolik¹³ method shows that these reflectances correspond to a value of ω_0 in the range 0.5–0.55, but with an uncertainty of ±0.1 because of assumptions made about the scattering properties of oil smoke. Second, we measured volume absorption coefficients from filters exposed in the plume using the integrating sandwich technique¹⁴, and these, when combined with the nephelometer scattering coefficients, give values for ω_0 in the range 0.55–0.70 (±0.1). Thus the likely range of ω_0 is 0.50–0.65, which is higher than previously published values for oil smoke^{4,7,10} but lower than that of 0.8–1.0 measured in non-carbonaceous aerosols¹³.

The broadband optical depth of the smoke was estimated from the change in downwelling shortwave flux as the aircraft ascended or descended through the plume. Figure 6 shows the

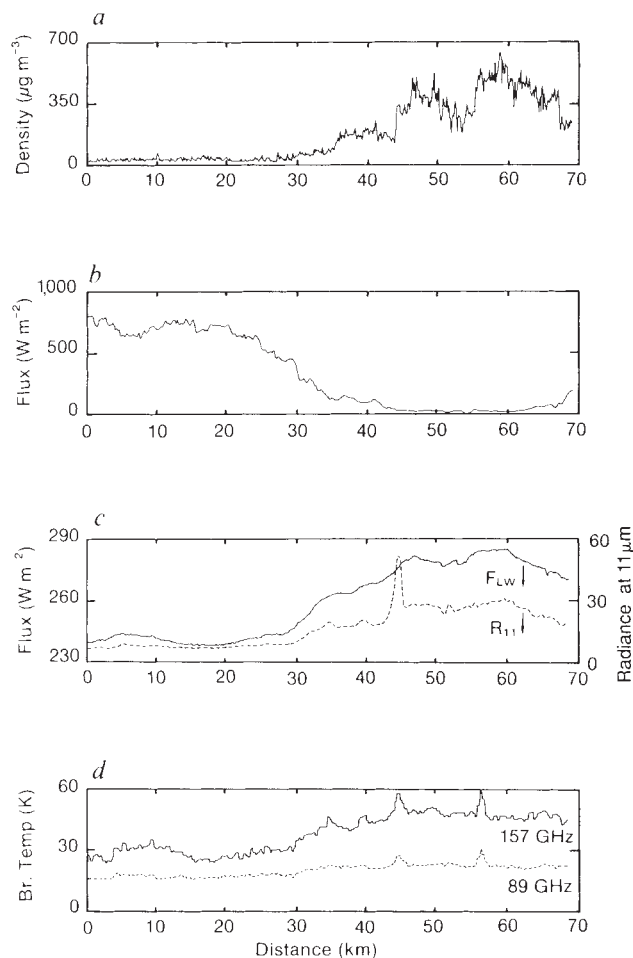


FIG. 5 Illustration of the effect of the smoke on downwelling radiation during a crosswind run through the centre of the plume. *a*, Smoke density; *b*, shortwave flux; *c*, infrared flux, F_{LW} , and radiance at 11 µm, R_{11} (in units of mW m⁻² steradians⁻¹ (cm⁻¹)⁻¹); *d*, the zenith microwave brightness temperatures at 89 and 157 GHz.

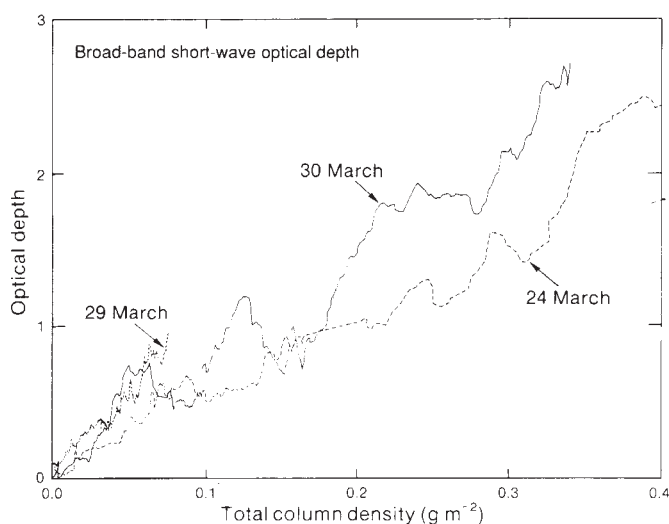


FIG. 6 Broad-band shortwave optical depths of the plume from three different ascending or descending profiles flown through the plume as a function of the total column density of smoke in the path. The profile on 29 March 1991 was further away from the source and lower smoke densities were encountered.

derived optical depth as a function of smoke column density for three profiles through the plume top. Solar heating rates in the plume were calculated from shortwave flux divergences in a profile at local noon on 30 March 1991, 140 km from the source. The heating rate was typically $\sim 50 \text{ K d}^{-1}$ throughout the depth (2,300 m) of the smoke. It was nearly constant in the plume because the smoke density increased from the top to the base. At a distance of 250 km from the source, the smoke density was a factor of four lower and the heating rates were 20 K d^{-1} . This compares with typical clear-sky values of 1 K d^{-1} .

Figure 5c shows the downward hemispheric infrared (4–50 μm) fluxes to be 45 W m^{-2} greater in the centre of the plume than in clear air and the vertical downward radiance at 11.1 μm to be increased by $20 \text{ mW m}^{-2} \text{ steradians}^{-1} (\text{cm}^{-1})^{-1}$. These were measured by a pyrgeometer¹⁵ and a narrow field of view (1.5°) radiometer¹⁶ respectively. The peak in radiance at 45 km along the run is probably due to cloud above and is also seen in the microwave measurements (Fig. 5d). The broadband infrared optical depth for a vertical path through the smoke was calculated in a similar manner to the visible optical depth. It was estimated to be 0.34 ± 0.05 for the profile on 30 March 1991. The 11.1- μm radiometer was not operated in this profile, but the maximum vertical optical depth calculated from the horizontal run in Fig. 5c was 0.30 ± 0.02 , which was 75% of the corresponding broadband optical depth. The infrared cooling rate in the plume was calculated from the flux divergences to be 2.1 K d^{-1} , which is 1.0 K d^{-1} greater than that in adjacent clear air. Assuming smoke densities based on the PCASP measure-

ments, we used the optical depth to deduce a visible extinction coefficient of $7\text{--}13 \text{ m}^2 \text{ g}^{-1}$, which agrees broadly with other measurements of oil smoke^{7,17}. The 11.1- μm infrared absorption coefficient was deduced to be $0.7\text{--}1.3 \text{ m}^2 \text{ g}^{-1}$, in agreement with laboratory measurements¹⁸.

Figure 5d shows the effect the plume had on downward radiances (or brightness temperatures) at 89 and 157 GHz, measured by a microwave radiometer. Atmospheric emission at these frequencies is principally controlled by water vapour and liquid water. The increase in brightness temperature in the plume at 157 GHz (22 K) is three times the increase at 89 GHz (8 K), but both are well correlated with the infrared radiance. This correlation was seen on several occasions when there was no trace of liquid water in the plume. The microwave radiance increase is surprisingly high and cannot be explained by variation in water vapour within the plume nor by the observed levels of SO_2 . Measurements¹⁹ of the absorption coefficient of diesel soot at 94 GHz are over an order of magnitude too small to account for our results.

Discussion

Our best estimate of oil burning rate is greater than those previously assumed^{1–3}. The greater CO_2 emissions that this implies, however, represent only 3% of total annual fossil-fuel emissions and would still have a negligible effect on global climate through the greenhouse effect. The sulphur emissions calculated are considerably higher than those taken by Browning *et al.*², and although we would expect the acidity of rainfall episodes to be no higher than previously predicted, because it is limited by droplet saturation, these episodes will extend further downwind across southern Asia. As predicted², increased photochemical oxidant concentrations were observed both close to the source and further afield. But the emissions of NO_x are smaller than those previously assumed²; for this reason and because they are rapidly removed, they are unlikely to influence the global NO_x background or global tropospheric ozone.

The estimated smoke emission rate, even with our large uncertainties, is similar to the range used in previous assessments^{1–3}. Furthermore, the product of our derived visible extinction coefficient and smoke emission rate (the factor that determines the radiative impact of the plume for a given meteorological situation) is independent of the uncertainties in our estimate of smoke density and is similar to that inferred from Browning *et al.*². Previous assessments^{2,3} assume that the smoke did not scatter radiation; the scattering that we observed would be insufficient to alter their conclusions significantly. The observed infrared effects, although previously ignored, will only very slightly ameliorate the predicted daytime surface cooling. Our measurements thus confirm the assumptions used in recent assessments^{1–3}, and the observation that significant amounts of smoke were not present above 5,000 m bears out their predictions. This study provides strong support for their conclusions that, although the effects may be significant on a regional scale, those on a global scale, including the Asian summer monsoon, are likely to be insignificant. □

Received 14 June; accepted 17 September 1991.

- Small, R. D. *Nature* **350**, 11–12 (1991).
- Browning, K. A. *et al.* *Nature* **351**, 363–367 (1991).
- Bakan, S. *et al.* *Nature* **351**, 367–371 (1991).
- Radke, L. F., Lyons, J. H., Hobbs, P. V. & Weiss, R. E. *J. geophys. Res.* **95**, 14071–14076 (1990).
- Marsh, A. R. W., Crabtree, J. & Hatton, D. B. *CERL RD/LI/2092N81* (Central Electricity Research Laboratories, Leatherhead, 1982).
- Crutzen, P. J. *Tropospheric Ozone*, 3–32 (Reidel, Dordrecht, 1988).
- Colbeck, I., Hardman, E. J. & Harrison, R. M. *J. Aerosol Sci.* **20**, 765–774 (1989).
- Colbeck, I., Appleby, L., Hardman, E. J. & Harrison, R. M. *J. Aerosol Sci.* **21**, 527–538 (1990).
- Berry, M. V. & Percival, I. C. *Opt. Acta* **33**, 577–591 (1986).
- Sokolik, I. N. *Izv. atmos. ocean. Phys.* **24**, 200–204 (1988).
- World Meteorological Organisation Rep. RDP 654: Meeting on atmospheric part of UN response to Kuwait oil-field fires, 27–30 April 1991; (WMO, Geneva, 1991).
- Environmental Impact of Oil burning in the Kuwaiti Oilfields: a Preliminary Evaluation* (King Fahd University of Petroleum and Minerals, Dhahran, 1991).

- Ginzburg, A. S. & Sokolik, I. N. *Izv. atmos. ocean. Phys.* **25**, 700–704 (1989).
- Kilsby, C. G. Q. *J. R. met. Soc.* **116**, 1173–1192 (1990).
- Foot, J. S. *J. Atmos. ocean. Technol.* **3**, 363–370 (1986).
- Rawlins, F. & Foot, J. S. *J. Atmos. Sci.* **47**, 2488–2503 (1990).
- Golitsyn, G. S., Shukurov, A. Kh., Ginzburg, A. S., Sutugin, A. G. & Andronova, A. V. *Izv. atmos. ocean. Phys.* **24**, 163–168 (1988).
- Roessler, D. M. & Faxvog, F. R. *J. opt. Soc. Am.* **69**, 1699–1704 (1979).
- Bruce, C. W., Stromberg, T. F., Gurton, K. F. & Mozer, J. B. *Appl. Opt.* **30**, 1537–1546 (1991).

ACKNOWLEDGEMENTS. We thank the Royal Air Force aircrew and ground crew; RAF Muharraq for facilities; the Mobile Met. Unit for forecasts; T. Marsh and P. Anwyl of National Power Technology and Environment Centre for the use and calibration of chemical monitors; P. Bridges of RAF Institute of Aviation Medicine; C. Gilmore of RAE Materials Department for SEM photographs, RAE Aircraft Department and RAF Lyneham for aircraft maintenance, and the scientists and technicians at the Meteorological Office for their hard work and commitment. The co-operation of Sheikh Essa bin Abdulla Al Khalifa and the State of Bahrain Ministry of Development and Industry, and that of the Gulf states who granted permission for the aircraft to operate in the area, was essential.

Sequence analysis of peptides bound to MHC class II molecules

Alexander Yu Rudensky, Paula Preston-Hurlburt, Soon-Cheol Hong, Avlin Barlow & Charles A. Janeway Jr*

Section of Immunobiology and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA

CD4 T cells recognize peptide fragments of foreign proteins bound to self class II molecules of the major histocompatibility complex (MHC). Naturally processed peptide fragments bound to MHC class II molecules are peptides of 13–17 amino acids which appear to be precessively truncated from the carboxy terminus, perhaps after binding to the MHC class II molecule. The finding of predominant self peptides has interesting implications for antigen processing and self–non-self discrimination.

ANTIGEN recognition by both antibodies and T-cell receptors is now fairly well understood¹. For antibodies, precise information is available in the form of X-ray crystallographic analysis of antigen–antibody complexes². For T-cell receptors, we know the primary structure of the receptor and we understand the generation of diversity that allows T cells to discriminate thousands or even millions of different proteins, but we lack a detailed picture of the precise mechanism of antigen recognition by T cells, due in part to the highly complex nature of the ligand. T-cell receptors do not bind antigen directly; rather, they recognize foreign antigens in the form of peptide fragments bound to a molecule encoded in the MHC³. Two MHC molecules have been characterized by X-ray crystallographic analysis^{4–6}: on the basis of these structures, all MHC molecules are believed to have a large peptide-binding cleft in their outer aspect, formed by two α helices overlying an eight-strand β -pleated sheet structure. The isolated and crystallized MHC molecules have peptides bound in this site⁵.

There are two classes of MHC molecule that derive peptides from two intracellular compartments^{7,8}. Peptides derived from proteins degraded in the cytosol and transported into the endoplasmic reticulum are an essential component of MHC class I molecules⁹. Peptides bound to MHC class I molecules have been characterized either by direct sequencing or by analogy to synthetic peptides^{10–14}. These peptides are 8–9 amino acids long^{11–14}; peptides of precisely this length are known to bind preferentially to MHC class I molecules¹⁵. The nature of the components of this complex are therefore well characterized.

MHC class II molecules, by contrast, derive their peptides from proteins that are degraded in cellular vesicles. These proteins enter the vesicular compartment by various routes, including phagocytosis, endocytosis, and internalization of self membrane molecules. The complex of peptide and MHC class II molecules is recognized by CD4 T cells¹⁶. As these cells are critical to most immune responses, the nature of the peptide–MHC class II complex is of central importance in immunology. To date, naturally occurring complexes have been identified and isolated^{17,18}. It is known that synthetic peptides can bind to isolated MHC class II molecules¹⁹ and that such complexes can be recognized by T cells²⁰. The length of peptides used in such analyses has varied widely. Information about naturally

processed peptides has not so far been available.

We have now purified MHC class II molecules and determined the sequence of peptides naturally associated with them. We observe that MHC class II-associated peptides are readily identified and characterized. Most derive from cellular proteins (self) or from proteins in the incubation medium (non-self), are 13–17 amino acids long and consist of relatively few distinct species. The detailed analysis of MHC class II-associated peptides suggests features of their generation from proteins that are distinct from those of MHC class I-associated peptides, perhaps reflecting differences in the structure of MHC class I and MHC class II molecules and the enzymes involved in their cleavage. Finally, our results have interesting implications for self tolerance.

Isolation of class II-associated peptides

To isolate MHC class II-associated peptides, we grew cells of the murine B-cell lymphoma line LB27.4 (H-2^{bxd}) (ref. 21) to high density in roller bottle cultures in medium supplemented with calf serum. Cells were collected, lysed, and the MHC class II molecules isolated by affinity purification over columns of anti-I-A^b (Y-3P) (ref. 22) or anti-I-E^b (Y-17) (ref. 23). Proteins purified according to ref. 24 are shown in Fig. 1a. They were eluted with¹⁷ acetic acid and any material of relative molecular mass (M_r) < 10,000 (10 K) was collected by filtration over a Centricon 10 membrane. This low-molecular-weight material was then fractionated by reversed-phase high-performance liquid chromatography (RP-HPLC) on a C18 column. Traces from I-A^b and I-E^b are shown in Fig. 1b and compared with a mock eluate from an equal amount of bovine serum albumin (BSA). Six prominent peaks from each eluate were sequenced. Note that the pattern of peaks generated by material eluted from I-A^b molecules is clearly distinguishable from that from I-E^b molecules; the control eluate contains only nonspecific peaks. Also, there are a limited number of major peaks. These data suggest that individual MHC class II molecules bind distinct sets of major peptides.

Sequences of I-A^b and I-E^b-associated peptides

The peaks marked 1–6 for both I-A^b and I-E^b were sequenced, in some cases several times, to determine the structure and length of these peptides. Database searches were then conducted to determine their protein of origin. The results are shown in Table 1. Of the six peaks sequenced from I-A^b molecules, five yielded sequences that matched 100% with regions of proteins in the data bank. One peak yielded what appears to be a mixture of two peptides; a specific donor protein has not been identified. Of the six peaks sequenced from I-E^b molecules, five yielded peptides that matched 100% with proteins in the data bank. One peak yielded weak signals and the protein of origin could not be identified. Thus, the peaks obtained by RP-HPLC fractionation of MHC class II-associated material yield single definable peptide sequences in most cases and these can be assigned to proteins of known sequence. Furthermore, all these proteins are probably present in the vesicular system of the cells used to isolate MHC class II molecules and so are good candidates to be peptide donors to MHC class II molecules in these cells. All of these peptides contain 13 amino acids or more. Integration

* To whom correspondence should be addressed.

TABLE 1 Sequences of MHC class II-associated peptides

MHC molecule	Peptide donor	Peptide sequence	Position	Length	HPLC fraction
I-A ^b	MuLV env protein	HNEGFYVCPGPHRP	145-158	14	1
		-----	145-158	14	4
		-----	145-157	13	2
	I-E α chain	ASFEAQGALANIAVDKA	56-73	17	5
	Ii	KPVSQMRMATPLLMR	39-53	15	6
	Undefined	XYNAYNATPATLAVD		15	3
I-E ^b	MuLV env protein	SPSYVYHQFERRAKYK	454-469	16	2
		-----	454-469	16	3
		-----	454-468	15	4
		-----	454-467	14	1
	Bovine serum albumin	GKYLEIARRHPY(F)	141-153 (4)	13 (14)	6
	Undefined	XPQSYLIHXXXXIS		14	5

Amino-acid sequences are shown of the dominant I-A^b and I-E^b-associated peptides. The peaks numbered sequentially in Fig. 1 are arranged by sequence. The identity of the murine leukaemia virus (MuLV) from which these peptides derive is unknown; it could be either RadLV or AKV, or a recombinant. The amino-terminal histidine of the MuLV envelope peptide marked 145-158 (RadLV numbering) was the most likely candidate for this position, but was not confirmed. The phenylalanine in the I-E-associated BSA peptide was a weak signal, so is shown in parentheses. For precise definition of carboxy termini, see Table 2. The main peptide peaks derived from I-A^b and I-E^b were sequenced on an Applied Biosystem 477 gas-phase protein sequencer. Amino-acid sequences were searched for identity with proteins in the GENBANK database using the FASTA program⁴⁶. Only instances of 100% identity with a known sequence are attributed.

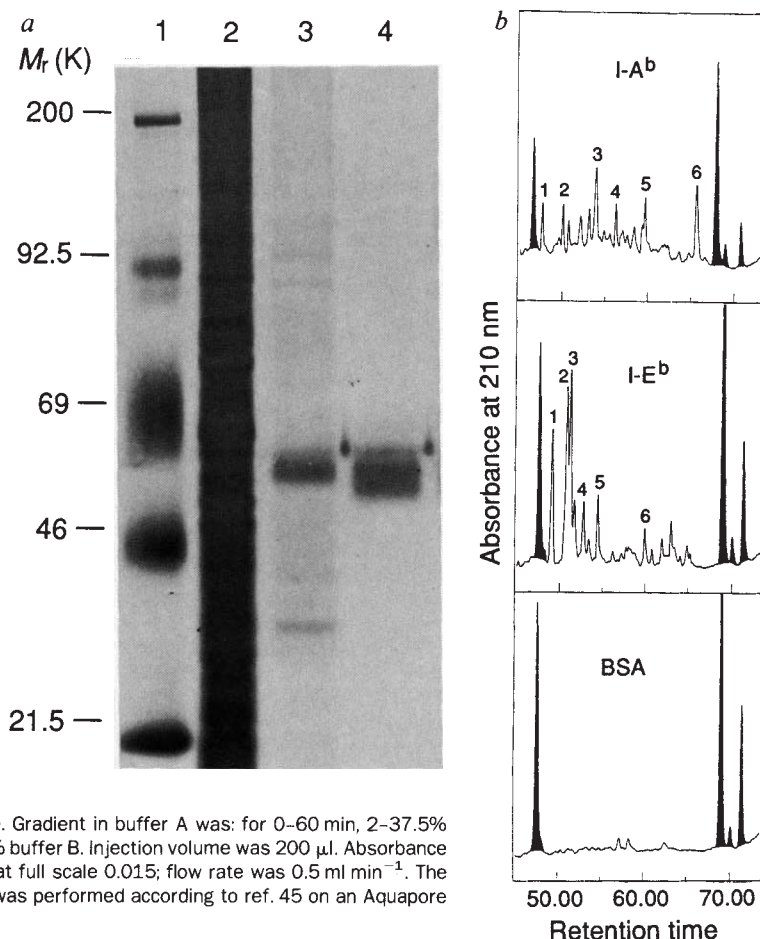
of the RP-HPLC traces shows that the six peptides analysed from I-A^b represent 51% of the total optical density units eluted in this region, and the six peaks analysed from I-E^b represent 74% of the eluted absorbance. Thus, they are a representative sample of the MHC class II-associated peptides. The peptides, the cleavages that generate them, and the precision with which they can be determined will now be discussed.

Peptides from transmembrane glycoproteins

Two of the peptides associated with I-A^b molecules derive from the extracellular domains of cellular transmembrane glycoproteins. I-A^b peak 6 is particularly prominent; it contains a peptide derived from the MHC class II-associated invariant chain (Ii) at positions 39-53. Ii binds to the chains of the MHC class II molecule in the endoplasmic reticulum, and remains

FIG. 1 Purified I-A^b and I-E^b molecules (a) and HPLC profiles of peptides eluted from these molecules (b). a, Silver-stained 10% SDS-polyacrylamide gel run under non-reducing conditions. Lane 1, standard proteins with M_r indicated in thousands on the left; lane 2, unseparated lysate; lane 3, affinity-purified I-E^b (3 μ g); lane 4, affinity-purified I-A^b (3 μ g). b, C18-column HPLC separations of peptides acid-eluted from the purified I-A^b and I-E^b shown in a, as well as control eluates from an equal amount of bovine serum albumin (BSA). Shaded peaks are seen in all three eluates and are an artefact; the numbered peaks were used for sequence analysis (Table 1). Peaks 2 and 3 from I-E^b were repurified by C8-column separation before analysis by sequencing (not shown).

METHODS. 1.8×10^{10} LB27.4 B lymphoma cells²¹ were grown in EHAA medium supplemented with 5% heat-inactivated calf serum and 50 μ g ml⁻¹ gentamycin. Cells were lysed at 2×10^8 cells ml⁻¹ of lysis buffer (containing PBS, 1% Nonidet-P40 detergent and a cocktail of protease inhibitors). I-A^b and I-E^b molecules were purified from cell lysate by affinity chromatography on Y-3P anti-I-A^b (ref. 22) and on Y-17 anti-I-E^b (ref. 23) monoclonal antibody columns as described²⁴ and concentrated on a Minicon B15 (Amicon). Protein concentration was determined by BCA assay (Pierce). Purification was monitored by sandwich enzyme-linked immunosorbent assay (ELISA) for I-A^b and I-E^b (S. Rath, R.-H. Lin, A.Y.R. and C.A.J., manuscript in preparation). Peptides were acid-eluted from 400-500 μ g of acetonitrile-precipitated I-A^b, I-E^b and BSA (negative control) as described¹⁷ and a low-molecular-weight fraction was isolated by centrifugation over Centricon 10 (Amicon) filters (~10K cutoff). This putative peptide fraction was fractionated by reverse-phase HPLC on a Vydac C18 column (250 $\times$ 4.6 mm 300 Å; 5 μ m) using a Waters HPLC system. Elution buffers: buffer A is 0.06% trifluoroacetic acid in water; buffer B is 0.052% trifluoroacetic acid in 80% acetonitrile. Gradient in buffer A was: for 0-60 min, 2-37.5% buffer B; for 60-90 min, 37.5-75% buffer B; for 90-105 min, 75-98% buffer B. Injection volume was 200 μ l. Absorbance of fractions was measured at 210 nm, shown as absorbance units at full scale 0.015; flow rate was 0.5 ml min⁻¹. The material in the major peaks marked was sequenced. Repurification was performed according to ref. 45 on an Aquapore C8 reverse-phase column (300 Å, 7 μ m); flow rate 0.15 ml min⁻¹.



associated with these molecules until they reach a *trans*-Golgi compartment where Ii is cleaved at several sites and dissociates from the MHC class II molecule²⁵. Ii-associated MHC class II molecules cannot bind peptides, indicating that Ii somehow occludes the peptide-binding cleft, preventing MHC class II proteins from transporting peptides from the endoplasmic reticulum in most cases^{26,27}. One possible explanation for finding this peptide associated with I-A^b is that it is a remnant of Ii that binds to a site unrelated to the peptide-binding groove. As it seems likely that Ii does bind MHC class II molecules at sites distinct from the groove, it is possible that such fragments could be associated with MHC class II molecules after cleavage. Alternatively, the peptide found could be the peptide of Ii that normally occludes the peptide binding groove; in either case, the same peptide should be found associated with a fraction of all MHC class II molecules. But we do not detect this peptide associated with I-E^b (Fig. 1 and Table 1): therefore, a third possibility is that this peptide is processed from Ii and binds specifically to the I-A^b groove like any other MHC class II-associated peptide; we favour this last possibility, but to prove it will require studies of several other *I-A* alleles.

A second peptide isolated in peak 5 derives from the α chain of the MHC class II protein I-E^b. We expected to find an E α peptide because I-A molecules can present processed I-E molecules²⁸. Furthermore, this E α peptide bound to I-A^b is recognized by a unique MHC-restricted monoclonal antibody, Y-Ae (ref. 29). Y-Ae binds I-A^b only on cells from mice that also express E α genes. T cells from C57BL/6J mice, which express I-A^b but not I-E, respond to this peptide presented by I-A^b (ref. 30). Thus, the E α peptide binds to the groove of I-A^b. The implications of these findings are discussed in more detail in the accompanying paper³⁰.

Peptides of exogenous proteins

MHC class II molecules present peptide fragments of foreign as well as self proteins derived either from extracellular spaces or from infectious agents residing in cellular vesicles. The peptide isolated as I-E^b peak 5 comes from bovine serum albumin in the serum in which the cells were cultured. This peptide is very similar in length to the other peptides identified. We have synthesized this peptide and have been able to show that B10.A(5R) mice (I-A^b, I-E^b) generate I-E-restricted T cell responses to it (Fig. 2a). Furthermore, this peptide inhibits responses of a T-cell hybrid to its peptide presented by I-E^b, whereas I-A^b binding peptides do not (Fig. 2b). Thus, this peptide clearly represents processed exogenous antigen bound to I-E^b.

Peptides of a retrovirus

Seven of the ten peptides identified derive from murine leukaemia virus (MuLV) envelope gp70 protein. These consist of two nested sets of peptides, one set associated with I-A^b and the other with I-E^b. The set associated with I-A^b completely matches gp70 of RadLV and differs by one amino acid from the same site in AKV. The set associated with I-E^b is 100% identical to AKV and differs by one amino acid from RadLV. The gp70 protein is a peripheral membrane protein and would thus be in the correct cellular compartment for processing. Furthermore, polymerase chain reaction (PCR) shows that a gene that could encode such a gp70 is expressed in the B lymphoma cells used to prepare the MHC class II molecules (Fig. 3). We cannot say whether both of the closely related AKV and RadLV are present in the cells, or whether a single virus with a point mutation or a recombinant gp70 gene is present that generates both peptides. Restriction enzyme digestion of

FIG. 2 The I-E^b binding peptide BSA 141–154 stimulates H-2^b lymph node cells from mice primed with BSA and is immunogenic for a recall response to BSA (a). Data are presented as Δ c.p.m.; $\blacksquare$, BSA peptide; $\square$, BSA. The optimal dose of BSA is >1 mg ml⁻¹, the amount used in culture of LB27.4 cells. b, The BSA 141–154 peptide ($\blacktriangle$) can competitively inhibit activation of the 2B4 I-E^b-restricted T-cell hybrid⁴⁷ to moth cytochrome c peptide 81–103 used at 10 μ g ml⁻¹ presented by B10.A (5R) (I-A^b, I-E^b) spleen cells. The I-A^b-binding peptide ($\bullet$) from *Staphylococcus* nuclease, residues 93–110 (pNase) (ref. 48), and the Ii 39–53 peptide ($\blacksquare$) found here associated with I-A^b, fail to inhibit this response.

METHODS. B10.A(5R) mice were immunized with BSA or BSA 141–154 peptide at 50 μ g per mouse in 100 μ l complete Freund's adjuvant (CFA). Lymph nodes harvested 8 days later were cultured at 4×10^5 per well in Click's EHAA medium with 1% normal mouse serum for 96 h at 37 °C in 5% CO₂ in air. Cultures were pulsed with 1 μ Ci ³H-thymidine for 16 h and harvested. Means of triplicate cultures are shown. Background c.p.m. in the absence of antigen subtracted from each group were 6,000 c.p.m. for peptide-primed mice and 5,000 c.p.m. for BSA-primed mice. The 2B4 hybrid was a gift from L. Samelson. Moth cytochrome c peptide 81–103 was provided by D. Levin and K. Bottomly. 2B4 cells (10^5 per well) were stimulated by irradiated B10.A(5R) spleen cells (2×10^5 per well) plus 10 μ g ml⁻¹ antigen. The indicated doses of competitor peptide were added simultaneously to the wells of 96-well plates. After 10 h interleukin-2 (IL-2) production was determined by a standard CTLL assay.

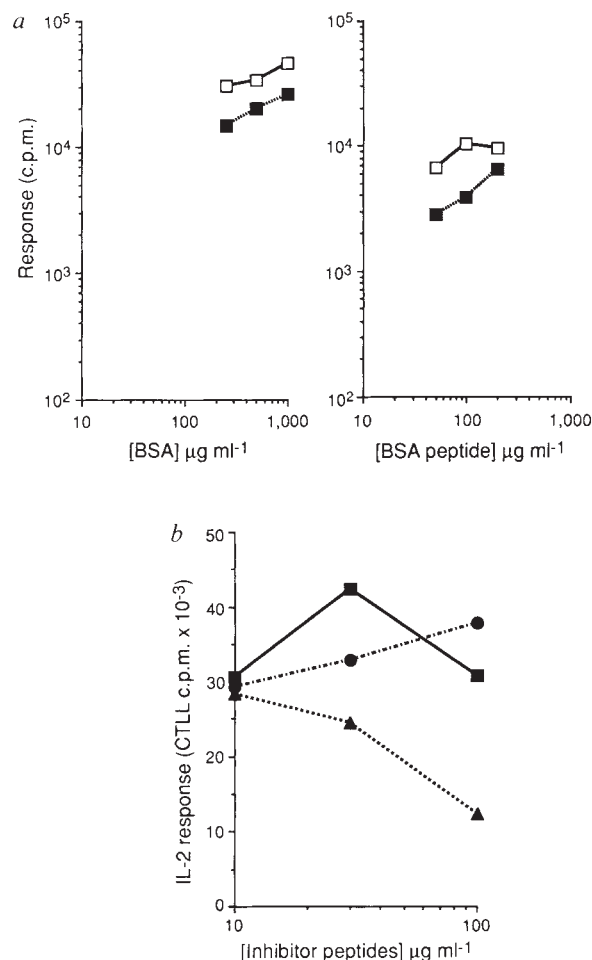


TABLE 2 Cleavage sites of MHC class II-associated peptides

MHC molecule	Peptide donor	N-terminus	C-terminus	Length	HPLC peak
I-A ^b	MuLV env protein	VTHA/ <u>HNEG</u>	PHRP/RWAR	14	1
			PHR/PRWA*	13	2
			VDKA/NLDV	17	5
I-E ^b	I-E α chain	FAKF/ <u>ASFE</u>	LLMR/PMSM*	15	6
	li	PKSA/KPVS	AKYK/REP*	16	2
	MuLV env	VTYH/ <u>SPSY</u>	AKY/KREP*	15	4
			AK/YKRE*	14	1
			RHPY/F/YAP	13 (14)	6
	Bovine serum albumin	KKFW/ <u>GKYL</u>			

Cleavage sites at the N and C termini of peptides derived from proteins of known sequence are indicated by /. The precise cleavage site is known for the N terminus in all five cases. Where the precise site of termination of the C-terminal sequence is known, based on predicted yield of the next amino acid from the sequencer, this is denoted by an asterisk. This identification is strengthened by finding the same peptide at different lengths in different fractions off the C18 column (see Fig. 1 and Table 1).

the PCR product in Fig. 3 is consistent with a single transcript indistinguishable from AKV gp70 (data not shown).

The finding of gp70-derived peptides is of interest because they represent processed viral peptides. Many immune responses to viruses need CD4 T cells, including antibody production and the generation of cytolytic T cells. The envelope proteins are the most likely viral component to be presented by MHC class II molecules in infected antigen-presenting cells, and these are evident here, contributing 7 of the 12 major peaks and several of the most prominent peaks (I-A^b, peaks 1, 2 and 4; and I-E^b, peaks 1-4). Other interesting features of this set of peptides, such as length differences and identification of what appears to be the same peptide in distinct peaks, are discussed below.

Peptides of unknown origin

Two of the peaks, peak 3 eluted from I-A^b and peak 6 eluted from I-E^b, contain peptides of sequences that cannot be matched to known proteins; these peaks may also contain more than one peptide, as good candidates for two different amino-acid residues are found at several positions. But they do contain peptides of roughly the same length as those deriving from known proteins. Finally, an unmarked peak eluted from I-A^b which was loaded into the sequencer yielded no amino acids; either it had a blocked amino terminus or it was not a peptide.

Features of class II-associated peptides

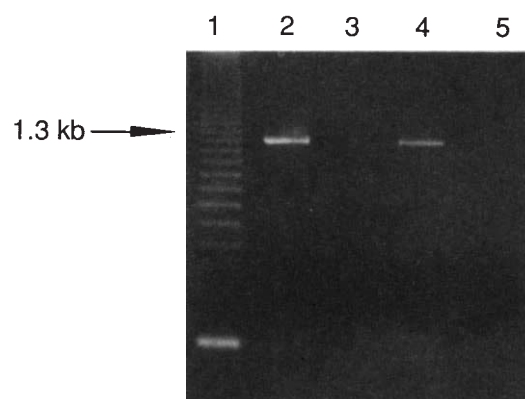
The finding that 10 of the peptides we sequenced were derived from proteins of known primary structure enabled us to examine the cleavage sites of these peptides for clues about the mechanism of processing. MHC class II-associated peptides are believed to be processed, in endosomal or lysosomal vesicles that require

acidification, by proteases inhibitable with leupeptin³¹. Cleavage of the invariant chain is inhibited by the same treatments, however, and may be an essential prerequisite for presentation^{25,32}. Thus, a direct examination of MHC class II-associated peptides may yield clues about the enzymes involved in the processing of MHC class II-associated peptides. As shown in Table 2, the five distinct amino termini of the ten peptides from identified proteins do not reveal a common cleavage site. Cathepsin B could be an enzyme involved in processing³³; this usually cleaves after a pair of basic amino acids, but only the sequence of the E α peptide contains a precise cathepsin B site. This indicates that distinct enzymes produce each different amino-terminal cleavage, or that another enzyme is involved (such as cathepsin D), or that cleavage by cathepsin B is followed by trimming by aminopeptidases (for example, the BSA peptide)³⁴.

In the case of the eight distinct carboxy termini of the ten peptides of known origin, we must first ask how precisely these are defined. As shown in Table 2, we would have expected to detect readily the next amino acid, based on repetitive yield and signal strength, in six out of the ten cases. In two other cases, there could be one additional amino acid, although we cannot detect it. In the case of the peptides of unknown origin, no comment can be made about the precise C terminus. Thus, the lengths shown in Table 1 are minimal length estimates in six cases and exact lengths in the other six, as noted. The most informative C termini come from the two sets of gp70-derived peptides. These differ from each other by one or two amino acids, the lengths are known with precision, and each was isolated as a discrete peak. This strongly suggests that the precise C terminus is determined by cleavage by a carboxypeptidase,

FIG. 3 Detection of messenger RNA encoding murine leukaemia virus, probably AKV, gp70 envelope protein by polymerase chain reaction (PCR) of LB27.4 complementary DNA. Shown are cDNA derived from LB27.4 (lane 2) and 3 control cells: 4G4, a hybrid with the AKR thymoma BW5147, containing AKV (lane 4); lipopolysaccharide-activated B cells of C57BL/6 mice (lane 3); lipopolysaccharide-activated B cells of (BALB/c $\times$ C57BL/6)_F₁ mice, identical in MHC to LB27.4 (lane 5). Lane 1, molecular size markers of 123-base pair multimers. A band at ~1.3 kilobases is generated by primers lying 5' of the env peptide associated with I-A^b and 3' of the env peptide associated with I-E^b. The PCR product is the same size as would be predicted for either AKV or RadLV.

METHODS. RNA was prepared from LB27.4 and control cells. First strand cDNA was synthesized by reverse transcriptase (Seigaku) using oligo(dT) as a primer and 1 μ g total RNA. PCR amplification was performed using a DNA thermocycler and *Taq* DNA polymerase (Perkin Elmer-Cetus) and reaction conditions recommended by the manufacturer (100 nM Tris, pH 8.3, 500 mM KCl, 15 mM MgCl₂, 0.1% gelatin). Thirty cycles were run for 1 min at 94 °C, 1 min at 55 °C, and 2 min at 72 °C. The sequence of the sense primer was CTCCACGCCAGGCTGTCCA and that of the anti-sense primer TCTAGGCCTACGATTCTG. Precipitated products were loaded onto agarose minigels with ethidium bromide for staining.



perhaps after the peptide has bound in the groove of an MHC class II molecule³⁴.

In two cases, peptides that are identical in sequence migrate as two different RP-HPLC peaks. The difference between the peptides associated with I-A^b peaks 1 and 4 must be dramatic, as these run with strikingly different mobilities on RP-HPLC. As the C termini of these two peptides is not known with certainty, they could differ in length. The I-E^b-associated peaks 2 and 3 were resolved by further C8 RP-HPLC (not shown). The peptides associated with I-E^b peaks 2 and 3 have the same sequence and length. The material in the two peaks could have been modified by oxidation, phosphorylation or deamidation, for example. Modifications could have occurred during purification. Alternatively, the changes may have happened in the cell; modification by live cells of a minimal peptide determinant bound to I-E^b has been reported which renders the peptide more immunogenic³⁵. Such peptides deserve further chemical analysis.

Length of class II-associated peptides

Sequence analysis of peptides isolated directly from MHC class I molecules¹¹⁻¹⁴, as well as analysis of synthetic peptide binding to isolated MHC class I molecules¹⁵, reveals that peptides of 8-9 amino acids are the natural form of MHC class I binding peptides. Recently, eleven peptides of this type have been isolated from purified HLA-B27 MHC class I molecules and sequenced¹⁴. Assuming that the peptides we have identified bind in the peptide-binding groove of MHC class II molecules, then such peptides are clearly distinct from those associated with MHC class I molecules, ranging from a minimum of 13 amino acids up to at least 17 amino acids. A larger size has been estimated for self MHC class II-associated peptides as well³⁶. Peptides of similar length are isolated from I-A^b and I-E^b. Although these peptides may be longer than estimated, they cannot be shorter. MHC class II molecules therefore not only allow longer peptides to bind than do MHC class I molecules, they seem to bind such peptides preferentially. This finding is in keeping with the observation that, whereas CD4 T cells can be stimulated by peptides of minimum length 8 amino acids longer peptides are usually more potent³⁷, perhaps reflecting better binding of the longer peptides to MHC class II molecules.

It has been proposed that the MHC class II molecule is similar in structure to the MHC class I molecule³⁸, specifically in the nature of the peptide-binding groove. If MHC class I molecules bind peptides of 8-9 amino acids, why should peptides binding to the MHC class II molecule be 13-17 amino acids long? One possibility is that one end of the peptide-binding groove of MHC class II molecules may be open, while both ends of the groove are closed in MHC class I molecules (refs 4-6; D. C. Wiley and P. J. Bjorkman, personal communication). One end of the groove on all MHC class I molecules is occluded by an invariant tyrosine, the other has an invariant salt bridge. MHC class II molecules may have a salt bridge at one end of the groove, but the other end has neither a potential salt bridge nor an invariant tyrosine or other bulky side chain. This may explain why longer peptides are preferentially presented by MHC class II molecules.

Finally, contrary to recent reports of sequence motifs in MHC class I-associated peptides^{13,14}, we do not find simple patterns of amino acids in the peptides isolated so far. This could reflect small sample size (four distinct peptides for I-A^b, and three for I-E^b), the differences allowed in MHC class II-associated peptide length, or a greater flexibility in peptide binding by MHC class II molecules.

Binding in the MHC class II groove

The peptides we have identified could bind to any site on the MHC class II molecule; they are interesting only if they bind to the MHC class II peptide-binding groove, which is the most likely explanation for our findings. Indeed, both the BSA peptide

and the Ea α peptide are immunogenic in appropriate strains of mice (Fig. 2; ref. 30), and the BSA peptide inhibits peptide binding to I-E^b but not I-A^b, whereas the Ii peptide blocks peptide binding to I-A^b but not I-E^b (Fig. 2b, and other data not shown). In addition, all the peptides are of a similar length, the binding to MHC class II molecules resists the purification procedure, the identified peptides are all derived from proteins in the expected cellular compartment, and the peptides isolated from different MHC class II molecules are different. The peptides identified here probably all bind in the peptide-binding groove, but eventually their ability to block the binding of known immunogenic peptides will be required to prove this point for each peptide.

Implications for processing

Our data suggest a model for the processing of peptides presented by MHC class II molecules that differs somewhat from the processing of peptides that bind MHC class I molecules proposed by Falk *et al.*³⁴. Our model is shown in schematic form in Fig. 4. MHC class I molecules are assumed to have a peptide binding groove with fixed dimensions that allows only 8-9 amino acid peptides to bind. MHC class II molecules would have one end of the peptide-binding groove open, and the N terminus of the peptide bound in the occluded end of the groove. The N terminus is generated independently of the MHC class II molecule, and the peptide binds in the groove with its C terminus protruding. Alternatively, both ends of the peptide could protrude from the groove after rough cleavage and binding, and amino peptidases could trim the N terminus until it fits the occluded end of the MHC class II groove exactly. The C terminus would then be cleaved by carboxypeptidases to achieve a final length that lies largely or entirely within the peptide binding groove. The variability seen at the C terminus would reflect this final trimming.

Our finding of several highly abundant peptides derived from endogenously synthesized membrane proteins bound to the peptide-binding site of MHC class II molecules raises other issues of importance in immunobiology. These endogenously synthesized peptides probably occupy most of the MHC class II molecules on the cell³⁶, undoubtedly affecting the recognition of foreign antigens³⁹, of non-self MHC molecules by T cells in the phenomenon known as alloreactivity⁴⁰⁻⁴², and the specificity of the T cell receptor repertoire^{29,43}. The ability of endogenous peptides to compete with exogenous antigens is often discussed as a potential mechanism to inhibit adaptive immune responses. But protective immunity is largely directed at particulate antigens that are readily internalized and would be expected to compete effectively with endogenous peptides for presentation

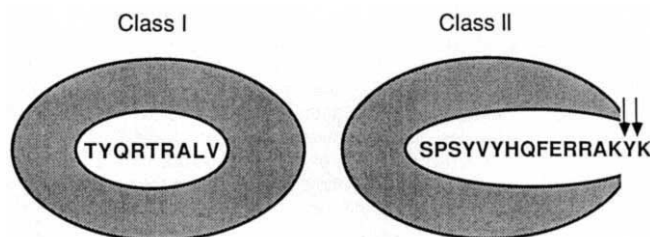


FIG. 4 A model for the observed differences between MHC class I- and MHC class II-associated peptides. It is known that the MHC class I site is closed at both ends, one end by a tyrosine side chain and the other by a salt bridge between residues on the two α helices. An MHC class I binding peptide is shown in this site. In MHC class II molecules, the tyrosine is replaced by a salt bridge formed by β -chain Asp 57 and α -chain Arg 88 in most alleles⁴⁹. It is predicted that the N terminus of the peptides detected here will bind in the closed end of the peptide-binding groove, and that the C terminus will extend out of the open end and be susceptible to trimming by carboxypeptidases, as noted by arrows above the sequence of the I-E^b binding MuLV envelope peptide 454-469.

by MHC class II molecules. The fact that we find peptides from a viral envelope protein and from BSA at 1 mg ml⁻¹ bound to MHC class II molecules further emphasizes the ability of abundant proteins or proteins of infectious agents to be presented.

Endogenous peptides may play a part in T-cell responses to non-self MHC class II molecules⁴², but they may be less heterogeneous than originally imagined. Five or so distinct peptides make up the 50–75% of the peptides associated with each of the MHC class II molecules examined (Fig. 1b; Table 1). Several of the 12 peaks examined have an identical sequence, meaning that they are probably not distinct from the perspective of a T cell³⁷. Thus, alloreactivity to MHC class II molecules may be less peptide dependent than alloreactivity to MHC class I molecules⁴⁴. The peptides identified here will be interesting to test in responses of alloreactive T cells. One, the E α peptide, is clearly recognized by I-A^b-restricted T cells³⁰; whether it is detected by I-A^b alloreactive T cells can now be determined. Furthermore, the BSA peptide we identified is reminiscent of earlier work⁴⁰ describing alloreactivity that needs a peptide of human serum albumin.

Finally, our finding that the repertoire of endogenously synthesized MHC class II-associated peptides is limited and dominated by a few highly abundant peptides has interesting implications for the selection of the T-cell receptor repertoire. We assume that the profiles we have obtained are typical of antigen-presenting cells *in vivo*, an assumption that is supported by finding these peptides immunogenic and by the Y-Ae data presented elsewhere³⁰. If this is true, then a few self peptides presented by MHC class II molecules may be critical in shaping the development of immature T cells. These few self peptide/self MHC class II complexes would presumably delete relatively few T cells, leaving most of the self MHC class II-restricted T cells able to respond to non-self peptides presented by self MHC class II molecules in the periphery. This argument is discussed

in detail in the accompanying paper³⁰, where quantitative estimates of the abundance and distribution of one such self peptide can be made with confidence.

Summary

We have isolated and sequenced 12 peptides associated with two MHC class II molecules purified from a B-cell lymphoma line. Ten of these 12 peptides could be identified as coming from proteins known to be present in the cellular processing compartment where association with MHC class II molecules is believed to occur. The peptides come mainly from membrane proteins, but one peptide was derived from an extrinsic or foreign protein. Two of two peptides tested are immunogenic. Different MHC class II molecules bind different peptides. These results strongly suggest that we have identified naturally processed peptides bound in the peptide-binding groove that is the site of antigen presentation by the MHC class II molecule. MHC class II-associated peptides identified here are all 13–17 amino acids in length, longer than previously identified MHC class I-binding peptides. The peptides have precise N-terminal cleavages but varying C-terminal cleavages, suggesting that the N terminus is protected in the peptide-binding groove of the MHC class II molecule, the C terminus being accessible and precessively truncated to a minimum of 13 amino acids. The exact proteases involved in this process cannot be inferred from the peptides isolated. Finally, the finding of a few highly abundant peptides associated with MHC class II molecules has important implications for the processing of foreign antigens, for the generation of responses to non-self MHC class II molecules, and for the problem of the selection of the T-cell repertoire. This technique may be a powerful approach to the identification of T-cell epitopes from infectious agents or autoantigens, as peptide identification requires no prior information about the specificity of such immune responses. □

Received 10 June; accepted 6 September 1991.

1. Davis, M. M. & Bjorkman, P. J. *Nature* **334**, 395–402 (1988).
2. Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. *Science* **233**, 747–753 (1986).
3. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51–177 (1979).
4. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
5. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
6. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692–696 (1989).
7. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903–921 (1986).
8. Germain, R. N. *Nature* **322**, 687–689 (1986).
9. Townsend, A. *et al. Nature* **340**, 443–448 (1989).
10. Rötzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283–287 (1990).
11. Van Bleek, G. M. & Nathanson, S. G. *Nature* **348**, 213–216 (1990).
12. Rötzschke, O. *et al. Nature* **348**, 253–255 (1990).
13. Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290–296 (1991).
14. Jardtzyk, T., Lane, W., Robinson, R. A., Madden, D. & Wiley, D. C. *Nature* **353**, 326–329 (1991).
15. Schumacher, T. N. M. *et al. Nature* **350**, 703–706 (1991).
16. Swain, S. L. *Immun. Rev.* **74**, 129–156 (1983).
17. Demotz, S., Grey, H. M., Appella, E. & Sette, A. *Nature* **342**, 682–684 (1989).
18. Srinivasan, H. & Pierce, S. K. *Proc. natn. Acad. Sci. U.S.A.* **87**, 919–922 (1990).
19. Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E. & Unanue, E. R. *Nature* **317**, 359–360 (1985).
20. Watts, T. H., Brian, A. A., Kappler, J. W., Marrack, P. & McConnell, H. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7564–7568 (1984).
21. Kappler, J., White, J., Wegmann, D., Mustain, E. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3604–3607 (1982).
22. Janeway, C. A. Jr *et al. J. Immun.* **132**, 662–669 (1984).
23. Lerner, E. A. *et al. J. exp. Med.* **152**, 1085–1101 (1980).
24. Buus, S., Sette, A., Colon, S. M., Jenis, D. M. & Grey, H. M. *Cell* **47**, 1071–1077 (1986).
25. Cresswell, P., Blum, J. S., Marks, M. S. & Roche, P. A. *Cold Spring Harbor Symp. quant. Biol.* **54**, 309–318 (1989).
26. Roche, P. A. & Cresswell, P. *Nature* **345**, 615–618 (1990).
27. Teyton, L. *et al. Nature* **348**, 39–44 (1990).
28. Briner, T. J., Guillet, J.-G., Lai, M.-Z., Smith, J. A. & Geffer, M. L. in *T-cell Receptors* (eds Davis, M. M. & Kappler, J. W.) 309–316 (1987).
29. Murphy, D. B. *et al. Nature* **338**, 765–767 (1989).
30. Rudensky, A. Y., Rath, S., Preston-Hurlburt, P., Murphy, D. B. & Janeway, C. A. Jr *Nature* **353**, 660–662 (1991).
31. Möller, G. *Immun. Rev.* **98**, 1–187 (1988).
32. Blum, J. S. & Cresswell, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3975–3980 (1988).
33. Takahashi, H., Cease, K. B. & Berzofsky, J. A. *J. Immun.* **142**, 2221–2229 (1989).
34. Falk, K., Rötzschke, O. & Rammensee, H.-G. *Nature* **348**, 248–251 (1990).
35. Fox, B. S., Carbone, F. R., Germain, R. N., Paterson, Y. & Schwartz, R. H. *Nature* **331**, 538–540 (1988).
36. Buus, S., Sette, A., Colon, S. M. & Grey, H. M. *Science* **242**, 1045–1047 (1988).
37. Cease, K. B., Berkower, I., York-Jolley, J. & Berzofsky, J. A. *J. exp. Med.* **164**, 1779–1784 (1986).
38. Brown, J. H. *et al. Nature* **332**, 845–850 (1988).
39. Harding, C. V. & Unanue, E. R. *Nature* **346**, 574–576 (1990).
40. Panina-Bordignon, P., Corradin, G., Roosnek, E., Sette, A. & Lanzavecchia, A. *Science* **252**, 1548–1550 (1991).
41. Benoist, C. & Mathis, D. *Curr. Biol.* **1**, 143–144 (1991).
42. Matzinger, P. & Bevan, M. J. *Cell. Immun.* **29**, 1–5 (1977).
43. vonBoehmer, H., Teh, H. S. & Kisielow, P. *Immun. Today* **10**, 61–66 (1989).
44. Heath, W. R., Hurd, M. E., Carbone, F. R. & Sherman, L. A. *Nature* **341**, 749–752 (1989).
45. Stone, K. L. & Williams, K. R. in *Macromolecular Sequencing and Synthesis: Selected Methods and Application* 7–24 (Liss, New York, 1988).
46. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
47. Samelson, L. E., Germain, R. N. & Schwartz, R. N. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6972–6976 (1983).
48. Finnegan, A. *et al. J. exp. Med.* **164**, 897–910 (1986).
49. Acha-Orbea, H. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2435–2439 (1987).

ACKNOWLEDGEMENTS. We thank S. Rath for initiating these studies; the Biopolymer Core Facility at the Howard Hughes Medical Institute at Yale University School of Medicine's Center for Molecular Medicine; K. Stone and K. Williams for HPLC and sequence analysis of peptides; D. Murphy for his advice, and S. Morin for typing the manuscript. This research was supported by the NIH and the Howard Hughes Medical Institute.

Activation of BPV-1 replication *in vitro* by the transcription factor E2

Liu Yang, Rong Li, Ian J. Mohr, Robin Clark* & Michael R. Botchan

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720, USA

Soluble extracts from uninfected murine cells supplemented with purified viral E1 and E2 proteins support the replication of exogenously added papilloma virus DNA. The E2 transactivator stimulates the binding of the E1 replication protein to the minimal origin of replication and activates DNA replication. These results support the concept that transcription factors have a direct role in the initiation of DNA replication in eukaryotes by participating in the assembly of a complex at the origin of replication.

DNA REPLICATION is likely to be regulated through various pathways. For example, post-translational modification of the proteins required for DNA synthesis is known to be pivotal¹⁻⁵. Although proteins that bind to origins of replication are known also to function in the control of transcription^{6,7}, the roles of transcription factors in regulating chromosomal replication are ambiguous. Numerous experiments have, however, shown that tissue-specific gene expression is correlated with early replication of the active gene and its flanking DNA, whereas the same gene when inactive in another tissue replicates late in the cell cycle⁸. This implies a link between transcription control and replication control.

Bovine papilloma virus type 1 (BPV-1) provides a model for exploring the roles of transcription factors in eukaryotic DNA replication. In transformed cells, the viral chromosome is maintained as a stable nuclear plasmid replicating in synchrony with the host DNA. Two viral proteins, E1 and E2, are both necessary and sufficient for replication⁹. E1 is a protein of relative molecular mass 68,000 (M_r 68K) and viral DNA with mutations in this ATP-binding protein cannot be maintained as nuclear plasmids¹⁰. Three related site-specific DNA-binding proteins are encoded by the E2 open reading frame (ORF)¹¹: a 48K transactivator, and two repressors, E2C and E8/E2, lacking the activation domain. The 48K transactivator binds to DNA as a dimer, and in combination with cellular factors, including SP-1¹², activates transcription from a number of viral promoters. The relative concentrations of the E2 family of proteins thus intricately regulate the transcriptional programme of the viral plasmids. Studies from our laboratory showed that the 48K E2 protein could form a tight complex with the E1 protein¹³. Partially purified E1 displayed a weak specific DNA-binding activity, and this activity was markedly stimulated by E2. To facilitate mechanistic studies and to ascertain whether E2 has a direct role in DNA replication, we developed a cell-free replication system.

Replication *in vitro* depends on viral proteins

Cell-free extracts from virus-transformed cells (ID13) did not support the *in vitro* replication of exogenously added BPV-1 DNA (not shown). It seemed possible that virus-encoded proteins might be limiting. Therefore, E1, E2 and the E1/E2 com-

plex were overexpressed in a baculovirus expression system and the proteins purified by immunoaffinity chromatography as previously described¹³ (Fig. 1a). When the purified E1/E2 complex was added to cell-free extracts from mouse ID13 or FM3A cells¹⁴, replication activity was observed. Figure 1b shows that the replication products of pKSO plasmid co-migrated with supercoiled (I) and nicked (II) pKSO markers only when the FM3A extracts were supplemented with the E1/E2 complex. In addition, a broad band of replication intermediates (RI) and high molecular weight forms were seen.

Initially, plasmids containing the upstream regulatory region (URR)—previously shown to contain the origin of replication¹⁵—were used as templates for replication. Pure form-I DNA template was used to minimize repair synthesis. Completely replicated form-I DNA increased with smaller template targets; thus smaller templates provide favourable substrates. Neither the plasmid containing the late region of BPV-1 DNA (p3M) nor the vector (pKS) directed DNA replication (Fig. 1b). Several experiments suggest that the heterogeneous material labelled RI in Fig. 1b is replication intermediates. For example, the heterogeneous material digested with single-cut restriction enzymes migrated more slowly than did open circle DNA. Also, upon double digestion with single-cut enzymes and *DpnI* (which cuts unreplicated DNA), the heterogeneous products migrated faster than the full-length linear DNA but slower than the largest *DpnI* fragment (data not shown). Furthermore, the time course presented in Fig. 2 is consistent with a precursor-product relationship between the RI and the forms I and II DNA. Finally, Fig. 1c shows that the replicated DNA migrating with the mobility of supercoiled plasmid is completely resistant to hydrolysis by *DpnI*.

Table 1 summarizes some of the essential requirements and characteristics of the *in vitro* papilloma virus replication system. The inhibition by aphidicolin suggests that one or more of the cellular DNA polymerases α , δ or ϵ ¹⁶ are involved in BPV-1 replication. Furthermore, the resistance to α -amanitin implies that transcription *per se* mediated by the E2 protein and RNA

TABLE 1 Requirements for BPV DNA replication *in vitro*

Conditions	Relative replication (%)
Complete	100
– Template DNA	0
– ATP	18
– CTP, UTP and GTP	71
– Phosphocreatine and creatine phosphokinase	22
+ Aphidicolin (10 $\mu\text{g ml}^{-1}$)	0
+ Aphidicolin (30 $\mu\text{g ml}^{-1}$)	0
+ α -Amanitin (100 $\mu\text{g ml}^{-1}$)	99
+ α -Amanitin (250 $\mu\text{g ml}^{-1}$)	81
+ Camptothecin and VM-26 (40 $\mu\text{g ml}^{-1}$ each)	3

The 'complete' system is the standard reaction mixture described in Fig. 1 and set as 100% for relative replication comparisons. The incorporation of ³²P was measured by scintillation counting in EcoLite (ICN Biochemicals). The actual count incorporated for the complete reaction was 40,000 c.p.m. (9.4 pmol). The count for the reaction containing no template DNA was 1,000 (0.2 pmol) and set to 0%.

* Cetus Corporation, Emeryville, California 94608, USA.

polymerase II is irrelevant. The blocking of *in vitro* replication by topoisomerase (types I and II) inhibitors suggests that the reaction requires unwinding of the DNA duplex.

The kinetics of incorporation shown in Fig. 2a are consistent with a multicomponent or multistep reaction. After a lag period of approximately 15–20 min, the rate of synthesis increases for about 1 h before reaching a plateau. At the plateau about 7 pmol deoxynucleoside triphosphates (dNTPs) were synthesized into DNA in a 25- μ l reaction. Similar reaction kinetics have been reported for the SV40 *in vitro* replication system^{5,17,18}. To determine the initiation site and the directionality of DNA replication, the products from various time points were analysed after digestion with *Dra*I and *Bst*XI. If replication initiates from within the BPV-1 sequences, fragment D should be labelled first. Subsequently, if replication proceeds bidirectionally, other fragments should become labelled in proportion to their molecular weight and position with respect to a unique start site. As shown in Fig. 2b, a symmetrical curve peaking at fragment D is observed. The curves do not flatten out completely with time, as replication intermediates predominate in the reaction, even after 2 h incubation (Fig. 2a).

The minimal origin of replication

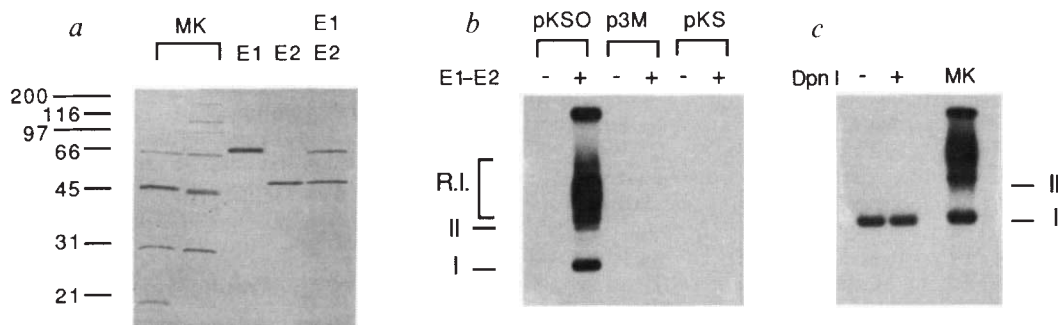
The series of plasmids used to localize the genetic elements necessary for BPV DNA replication *in vitro* are shown in Fig. 3. We were surprised to find that plasmid pC100 Δ RE replicated with the same efficiency as did the intact URR (Fig. 3), as this plasmid does not contain the highest affinity binding sites for

the E1/E2 complex¹³. At higher concentrations, however, the complex bound to other regions of the URR. In the absence of E2, E1 displayed a weak affinity for DNA proximal to but outside these high-affinity sites (see Fig. 3, and Mohr *et al.*¹³). Consistent with our data, *in vivo* studies of BPV-1 replication have supported the notion that these high-affinity sites are unnecessary *in cis* as genetic elements for replication⁹. Of the plasmids that replicated *in vitro*, pKSOM contains the least amount of viral DNA. This plasmid contains a part of E2 binding site 12 (ref. 12), an AT-rich region and an 18-base-pair (bp) palindromic sequence (Figs 3 and 4). This palindromic sequence motif is conserved in a number of animal and human papilloma viruses. To examine its genetic significance, mutants were created by inserting a synthetic linker into the *Hpa*I site (Fig. 4a). Neither pKSO-Nco nor pKSOM-Nco could support *in vitro* replication (Fig. 4b), suggesting that the spacing between palindromic half sites is important for replication *in vitro*.

E2 stimulates DNA replication

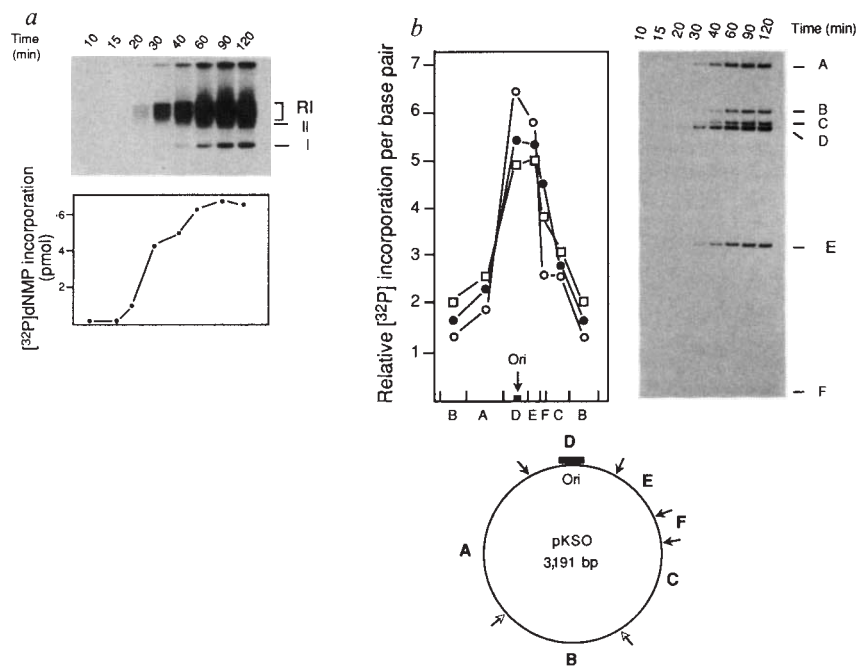
The *in vitro* replication system for BPV-1 was used to examine the role of E2 in DNA synthesis. Reactions receiving only E2 protein failed to replicate BPV-1 templates (Fig. 5, lanes 7 and 17). Reactions supplemented with purified E1 protein alone directed a small amount of replication at the highest levels of E1 only (Fig. 5, lanes 1 and 11). When purified E2 was added to extracts along with purified E1, replication was markedly stimulated (Fig. 5, top and bottom panels). A similar pattern of incorporation was detected with two different templates, pKSO

FIG. 1 BPV DNA replication *in vitro*. a, Purification of proteins. The BPV E1 and E2 coding sequences were cloned into a baculovirus expression system and the proteins were purified by immunoaffinity chromatography as described by Mohr *et al.*¹³. E1 was tagged at its N terminus with a peptide of 9 amino acids (EE epitope) and purified with a monoclonal antibody specific to the tag peptide (the antibody EE³¹ was crosslinked to protein G Sepharose, Pharmacia LKB). After washing the loaded column with LiCl, the protein was eluted with 20 mM triethylamine and concentrated (where necessary) by dialysis against solid polyethylene glycol (M_r 8,000), followed by extensive dialysis against 20 mM potassium phosphate (pH 7.5), 100 mM potassium glutamate, 1 mM EDTA, 1 mM DTT and 10% glycerol. In each lane 200 ng of the protein preparation was fractionated by SDS-PAGE and the proteins were stained with silver. The E2 and E1-E2 complex (designated E1/E2) were similarly purified with monoclonal antibody specific to E2 (B202)¹³. MK, molecular weight markers ($M_r \times 10^{-3}$). b, E1-E2 and BPV origin sequence-dependent *in vitro* DNA replication. In lanes labelled E1-E2+, 400 ng of the purified E1-E2 complex were added. Three template DNAs are shown here: pKSO contains BPV sequences from 7,805–100; p3M contains the BPV-1 restriction fragment *Hind*III (6,958)–*Mlu*I (7,351); and pKS is the vector plasmid (from Stratagene). I and II, form I and form II DNA; RI, replication intermediates. Both E1 and EE epitope at its N terminus and the wild-type E1 showed identical activities in ATPase and replication assays. However, E1 with EE epitope at its C terminus was inactive in both assays (data not shown). c, Replicated form I DNA is resistant to *Dpn*I digestion. The replication products and 200 ng of marker pKSO were mixed and separated by electrophoresis through a 1% SeaPlaque gel in 20 mM Tris-acetate (pH 8.0) buffer. The form I band was excised from the gel and hydrolysed with *Dpn*I. Ethidium bromide staining showed that all detectable form-I DNA was cleaved by the enzyme. MK, the pKSO replication products serve as a marker. METHODS. The conditions for BPV replication *in vitro* were derived from Li and Kelly³² with some modifications. Extracts from the mouse FM3A cell line were prepared as follows: cells were grown in a 2-l suspension culture containing RPMI 1640 media (supplemented with 25 mM HEPES, pH 7.2, and



5% calf serum). The cells were collected at a density of 7×10^5 cells ml^{-1} . The cell pellet was washed with 30 ml cold PBS and then 10 ml of hypotonic buffer (20 mM HEPES (pH 7.5), 5 mM KCl, 1 mM EDTA, 0.5 mM dithiothreitol (DTT)). The cells were then resuspended in hypotonic buffer to a final volume of 10 ml and incubated on ice for 15 min. After 20 strokes in a Dounce homogenizer (with B pestle), 500 μ l of 5 M NaCl was added and the extraction mixture was incubated on ice for 30–60 min. This mixture was centrifuged in a SW41 rotor at 20,000 r.p.m. for 30 min and the supernatant was dialysed twice against 1 l D buffer (20 mM HEPES, pH 7.5; 10 mM NaCl; 1 mM EDTA and 0.5 mM DTT). The extract was then centrifuged in HB4 rotor at 8,000 r.p.m. for 8 min and the supernatant was frozen as droplets into liquid nitrogen. The protein concentration of the extract was typically 15–20 mg ml^{-1} ; frozen extracts are kept at -70°C and are good for at least one year. A standard replication assay (25 μ l) contains: 10 μ l extract, 40–80 ng of pure form I template DNA, 30 mM HEPES (pH 7.5), 7 mM MgCl_2 , 20 mM potassium glutamate, 4 mM ATP, 100 μ M each of CTP, UTP and GTP, 26 μ M each of dATP, dTTP, dGTP and dCTP, 2.5 μ Ci each of the [^{32}P]dNTPs, 40 mM phosphocreatine and 100 μ g ml^{-1} creatine phosphokinase and viral proteins as indicated. The reaction was incubated at 37°C for 2 h, stopped by the addition of 25 μ l of 20 mM Tris (pH 7.7), 20 mM EDTA, 2% SDS and 50 μ g ml^{-1} proteinase K and incubated for another 30 min. The DNA was precipitated with 25 μ l of 7.5 M ammonium acetate and 175 μ l of 95% ethanol. The precipitation was repeated twice and the DNA was resuspended in 50 μ l Tris-EDTA. The DNA was analysed by electrophoresis in 0.8% agarose gel. Dried gels were exposed to X-ray film. Extracts from FM3A cells are capable of efficient repair synthesis. This activity can be measured with damaged DNA templates, is independent of virus-encoded proteins, and is essentially complete after 15 min of incubation (data not shown).

FIG. 2 BPV DNA replication *in vitro* initiates at an Ori-containing fragment and proceeds bidirectionally. *a*, Time course of BPV DNA replication *in vitro*. The replication assay was scaled up to 200 μ l. The reaction contained 640 ng of pKS0, 2.2 μ g of E1 and 0.6 μ g of E2. At each indicated time point, 25 μ l of reaction sample was taken and stopped. Top, an autoradiograph of the time course samples after electrophoresis. I and II, forms I and II of DNA. RI, replicating intermediates. Bottom, total incorporation into DNA of dNMP at each time point. *b*, Evidence for bidirectional replication. DNA samples from different time points were digested with *Dra*I and *Bst*XI, and the resulting seven DNA fragments (A–F) were separated in a 5% polyacrylamide gel. The autoradiogram of the gel is shown on the right. The intensity of each band was measured by the use of a Phosphor Imager (Molecular Dynamics). Incorporation per nucleotide was calculated for each fragment at each time point and the relative amounts of radioactivity were plotted (left side). $\circ$, 20 min; $\bullet$, 40 min; $\square$, 120 min. A diagram of the plasmid pKS0 is shown at the bottom. Open arrowheads, *Dra*I sites; solid arrowheads, *Bst*XI sites.



and pKSOM. The stimulation was due specifically to E2, as the E2C protein¹⁹, purified in a manner identical to E2, did not activate E1. The extent of E2 stimulation was dependent on the concentrations of both E2 and E1 (Fig. 5). At low concentrations of E1, replication was absolutely dependent on E2. The absolute levels of the E1 protein *in vivo* during S phase are not known, but we suspect that they are lower than those of the E2 protein, which we estimate at a few thousand molecules per cell. The absolute requirement for E2 *in vivo* may thus reflect, at least in part the low levels of E1 *in vivo*.

To determine whether interactions between E1 and E2 might mediate cooperative DNA binding, we carried out DNase footprinting studies. Figure 6a and b shows a DNase footprint analysis of purified E1 protein in the presence and absence of purified E2. E1 alone clearly protects DNA sequences centred over the 18-bp palindrome (labelled Ori) of pKS0. The linker insertion mutation pKS0-Nco dramatically diminishes this pro-

tection (data not shown). The E2 protein does indeed enhance the DNA binding capacity of the E1 protein. The concentration of E1 required to protect the E1 binding site to a certain level is 10-fold lower when E2 is present than in its absence. Cooperativity was, however, also seen with template pKSOM which lacks intact E2 binding sites. The ultraviolet-crosslinking experiment shown in Fig. 6c extends this observation. E2 could not be crosslinked to the DNA in the absence of E1, because no E2 sites exist in this target. However, in the presence of E1 the E2 protein can be crosslinked to the DNA (Fig. 6c, lanes C and D). Together with the footprint analysis, this clearly shows that the E1 and E2 proteins help stabilize the formation of a complex of both proteins over the replication origin.

E2 is a replication factor

The E2 proteins encoded by the papilloma viruses were first described and studied as transcription enhancer proteins²⁰. This

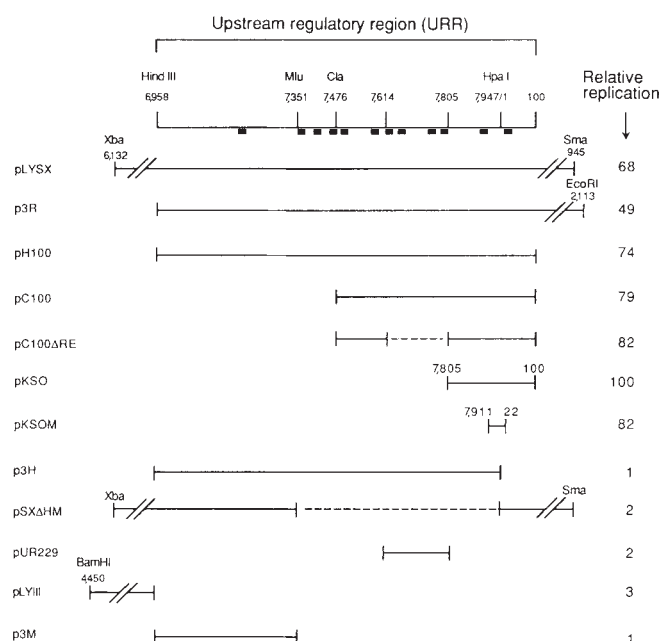
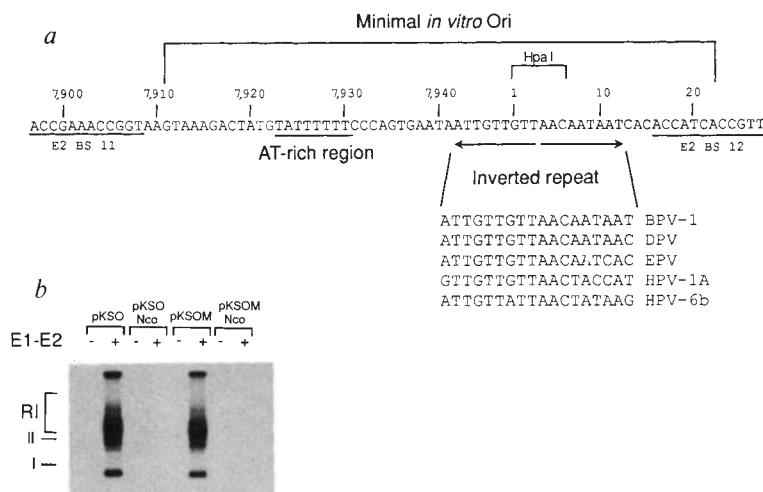


FIG. 3 Deletion analysis identifies *cis* elements required for replication. The physical map at the top shows the BPV URR which contains 12 E2 binding sites depicted as black boxes. DNA fragments spanning different regions of the viral genome were tested for their ability to function as an origin of DNA replication in the *in vitro* system. All reactions were carried out as described in Fig. 1 with 50 ng of DNA for each reaction. The results were quantified by two protocols: direct counting of the incorporated label; and by scanning and integrating each gel lane for each sample. 6 pmol of net synthesis were obtained with pKS0, and this number was set as 100%. No discrete bands were detected for templates p3H to p3M (see for example, Fig. 1b), and these templates are judged to be completely ineffective for *in vitro* DNA replication. pKSOM spans nucleotides 7911–22 (59 bp) and is the smallest fragment tested to date which shows replication activity. Where coordinate numbers are given, the polymerase chain reaction (PCR) was used to create the BPV insert placed into the plasmid polylinker. All other fragments were inserted into the polylinker of the vector by restriction site fusions. For both pKS0 and pKSOM, a *Bam*HI site was generated at the 5'-end and an *Eco*RI site at the 3'-end by PCR with primers.

FIG. 4 Structure of the *in vitro* origin of BPV DNA replication. *a*, The top line shows the BPV sequence from E2 binding site 11 (E2 BS 11) to E2 binding site 12 (E2 BS 12). An 18-bp inverted repeat centred at the *Hpa*I site is indicated by two tail-to-tail arrows. There is extensive homology between this inverted repeat and sequences present in the regulatory regions of the deer, elk and human papilloma viruses. *b*, Linker insertion at the centre of the inverted repeat abolishes replication *in vitro*. An *Nco*I linker (CATGCCATGGCATG) was inserted at the *Hpa*I site of pKSO and pKSOM. Standard replication assays were run on wild-type and mutant templates. The lanes labelled E1-E2+ contained 400 ng of the E1-E2 complex. All reactions were incubated at 37°C for 2 h.



study shows that E2 is a replication factor. The *in vitro* results show that E2 is not required to stimulate transcription of host genes encoding replication factors, nor to stimulate transcription of the template to create a primer. A possible role for E2 is that it helps target and stabilize a preinitiation complex at the origin site together with the E1 protein. By a combination of protein-protein and protein-DNA interactions, E2 may, for example, stabilize E1 in a conformation that facilitates origin recognition. The protein complex which assembles over the origin is quite large as measured by DNase I protection at high protein concentration, as the footprint extends in both directions past the 18-bp palindrome. Assembly of such large protein complexes which extend past the genetic determinants for initial binding may be a general feature of replication origins^{21,22}. Stabilizing the E1 protein to a site, however, is not sufficient for initiation. E1, either by itself or complexed with E2, also binds to a region just upstream of the origin between the high affinity E2 binding sites known as E2 RE1. The DNase I protection, centred around nucleotide 7,690, is similar to the protection found at the origin (data not shown). The significance of this particular interaction is unclear as these sites are not necessary for replication *in vitro* or *in vivo*, nor will they function as an origin of replication *in vitro*. Nevertheless, they may influence the actual position from which bidirectional replication initiates. Alternatively, they may be important for the regulation of transcription and replication. Presumably the E1 protein has some special function in replication which requires a structure dependent on the sequences present only at the origin. Preliminary data indicate that the E1 protein is a helicase (I.J.M. and M.B., unpublished observations).

Previous work in this laboratory localized the BPV-1 origin of replication to nucleotides 7,730 ± 100 bp by two-dimensional gel mapping¹⁵. Origins identified by two-dimensional gel electrophoresis reflect the position at which stable duplex formation first occurs on one or both sides of the replication intermediate. The replication intermediate may or may not include the genetic determinants that direct the assembly of the initiation complex. The limitations of two-dimensional gel mapping are revealed in our present study which localizes the genetic determinants for origin function *in vitro* to nucleotides 7911–22, approximately 100 bp downstream from the region identified by the two-dimensional gel method. The region defined as the genetic origin in this study also functions *in vivo* (M. Ustav and A. Stenlund, personal communication; R. Mendoza and M.B., unpublished observations). These *in vivo* studies also show that intact E2 binding sites are not required on the templates although E2 protein is required *in trans* for replication.

SV40 and polyoma large T antigens provided the first mechanistic insights into how cellular eukaryotic DNA poly-

merases are assembled at specific DNA sites to initiate replication. In contrast to these viruses, papillomaviruses regulate their DNA replication to maintain a constant copy number in dividing cells. Comparisons between the two systems may thus be informative. The detailed structure assembled at the BPV origin is unknown, and we do not know if the E1 and E2 proteins must assemble as independent units on the DNA, or if a preformed complex binds to the origin. Although the purified E1-E2 complex directs DNA synthesis when added to replication reactions, the complex may first dissociate and then reform on BPV origin DNA. Preformed SV40 T antigen hexamers do not initiate

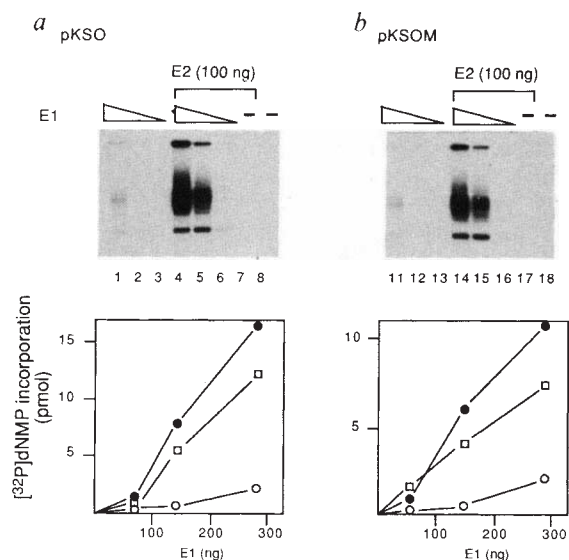


FIG. 5 E2 stimulates E1-dependent replication. *a*, pKSO which contains both E2 binding site 11 and 12 was used as the template DNA. Standard replication assays were performed either in the presence or absence of E1 and E2 proteins at 37°C for 2 h. Top, an autoradiogram of the replication products after fractionation by gel electrophoresis. The amount of viral proteins in each reaction is: E1, 280 ng (lanes 1 and 4); 140 ng (lanes 2 and 5); 70 ng (lanes 3 and 6). E2: 100 ng (lanes 4–7). Lane 8, control with no added protein. Bottom, a set of E1 and E2 titration experiments were carried out and the ³²P incorporation was measured and plotted as shown. *b*, pKSOM which contains no E2 binding site was used as the template DNA. Top, autoradiogram of the replication products. E1: 280 ng (lanes 11 and 14); 140 ng (lanes 12 and 15); 70 ng (lanes 13 and 16). E2: 100 ng (lanes 14–17). Lane 18, control with no added protein. Bottom, the titration of E1 and E2 protein concentrations and replication for the pKSOM template. ○, No E2; ●, 100 ng of E2; □, 30 ng of E2. E2 amounts greater than 100 ng gave no further stimulation (not shown).

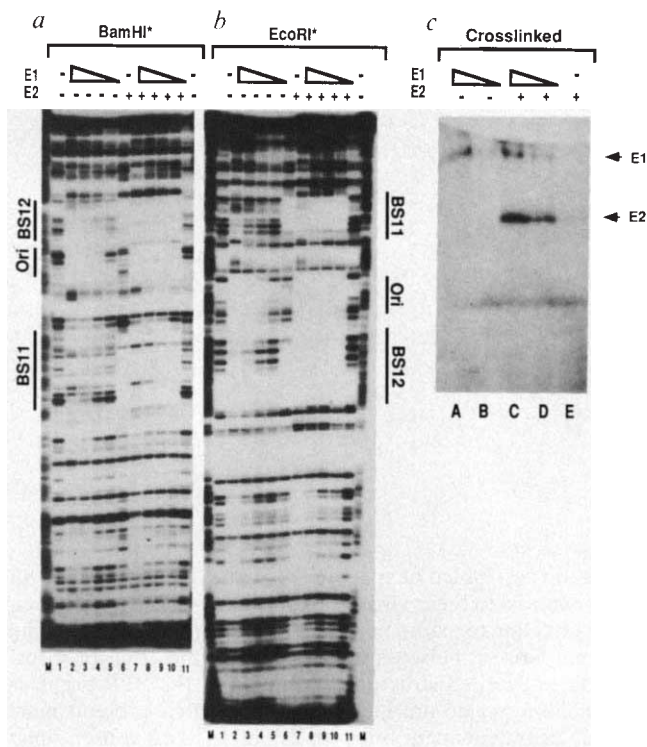


FIG. 6 *a* and *b*, Binding of E1 to the origin of BPV replication is stimulated by E2. DNA fragments containing BPV sequence 7,805–100 were labelled with ^{32}P at the 5'-end of each strand (*a*, top strand with labelling at *Bam*HI site; *b*, bottom strand with labelling at *Eco*RI site). The DNase I footprint analysis was performed as in ref. 12 with following modifications. The Z buffer was replaced with a new buffer containing 20 mM potassium phosphate (pH 7.5), 100 mM potassium glutamate, 1 mM EDTA, 0.5 mM DTT and 10% glycerol. The binding reaction was carried out at 37°C for 15 min, followed by standard DNase I digestion. E1 concentration: lanes 2 and 7, 900 ng; lanes 3 and 8, 300 ng; lanes 4 and 9, 100 ng; lanes 5 and 10, 33 ng. E2 concentration: lanes 6–10, 100 ng. BS11 and BS12 are E2 binding sites 11 and 12. M, AG sequence size marker. *c*, E1 is required for the interaction of E2 with DNA in the absence of E2 binding sites. The interaction of E1 and E2 with DNA were probed by UV-crosslinking of the proteins to a ^{32}P -labelled bromodeoxyuridine-substituted DNA containing the minimal replication origin (no E2 binding sites). The DNA was subsequently digested and the proteins were analysed by polyacrylamide gel electrophoresis. For E1 protein concentration: lanes A and C, 280 ng; lanes B and D, 90 ng. For E2 concentration: lanes C–E, 100 ng.

METHODS. The experimental protocol for UV-crosslinking was carried out as previously described³³. A primer annealed to the single-stranded pKSOM was extended by Klenow DNA polymerase in the presence of dCTP, dGTP, [α - ^{32}P]dATP and 5-bromo-2'-deoxyuridine triphosphate. The double-stranded DNA was digested with restriction enzymes *Bam*HI and *Eco*RI, and the DNA fragment containing the minimal replication origin was isolated and used in the crosslinking reaction. E1 and E2 proteins were incubated with the labelled DNA at 37°C for 30 min HEPES (pH 7.5), 7 mM MgCl_2 , and 100 mM potassium glutamate. The reaction mixtures were then irradiated with ultraviolet light for 60 min at room temperature. After digestion with DNase I and micrococcal nuclease, the E1, E2 proteins were separated in a 12% polyacrylamide gel by electrophoresis and detected by autoradiography.

replication²³ (F. Dean and J. Hurwitz, personal communication); however, hexamers are assembled on each half site of the palindromic DNA sequence that defines the SV40 origin and subsequently direct the initiation of DNA synthesis^{24,25}.

Cheng and Kelly²⁶ have shown that the transcription factor NF-1/CTF can overcome the inhibiting effects imposed by chromatin on *in vitro* SV40 replication, but only when NF-1 was prebound to the template before chromatin assembly. In the BPV *in vitro* replication system, we assume that chromatin is formed because completely replicated molecules migrate with the expected superhelical density. Furthermore, only relaxed DNA is detected in the absence of DNA replication when DNA templates are incubated with extract. Chromatin assembly is coupled to DNA replication in other systems and thus occurs after initiation^{17,18}. Therefore, we do not think it probable that chromatin assembly affects our present data, as only one round of replication is likely to be occurring *in vitro* (data not shown). E2, however, may stimulate replication through several direct

pathways, one of which may indeed perturb nucleosome distribution. In fact, it seems likely that site-specific DNA-binding transcription factors, when functioning as preinitiation complexes for RNA transcription, may have multiple roles, one of which is likely to be related to nucleosome positioning²⁷. Thus, the E2 protein may have multiple protein domains which function in different ways for replication.

The papilloma viruses were the first DNA tumour viruses described²⁸ and the human papilloma viruses are under intense study as the probable major aetiological agents in cervical cancer^{29,30}. The development of an *in vitro* DNA replication system for BPV-1 may prove helpful in dissecting cellular components involved in critical steps of the viral life cycle. Moreover, given the significant homologies in DNA sequence between the E1 binding site within the BPV-1 origin of replication and those in the upstream regulatory region of the human viruses, it is likely that the knowledge obtained with one system will be applicable to the other. □

Received 17 June; accepted 9 September 1991.

- Mohr, I. J., Stillman, B. & Gluzman, Y. *EMBO J.* **6**, 153–160 (1987).
- McVey, D. *et al. Nature* **341**, 503–507 (1989).
- D'Urso, G., Marraccino, R. L., Marshak, D. R. & Roberts, J. M. *Science* **250**, 786–791 (1990).
- Din, S. U., Brill, S. J., Fairman, M. P. & Stillman, B. *Genes Dev.* **4**, 968–977 (1990).
- Virshup, D. M., Kauffman, M. G. & Kelly, T. J. *EMBO J.* **8**, 3891–3818 (1989).
- Depamphilis, M. L. *Cell* **52**, 635 (1988).
- Brand, A. H., Micklem, G. & Nasmyth, K. *Cell* **51**, 709 (1987).
- Hatton, K. S. *et al. Cancer Cell* **6**, 335–340 (1988).
- Ustav, M. & Stenlund, A. *EMBO J.* **10**, 449–457 (1991).
- Sun, S., Thorner, L., Lentz, M., MacPherson, P. & Botchan, M. J. *Virology* **64**, 5093–5105 (1990).
- Howley, P. M. in *Virology* (eds Fields, B. N. & Knipe, D. M.) 1625–1650 (Raven, New York, 1990).
- Li, R., Knight, J. D., Jackson, S. P., Tjian, R. & Botchan, M. R. *Cell* **65**, 380–400 (1991).
- Mohr, I. J. *et al. Science* **250**, 1694–1699 (1990).
- Nakano, N. *J. exp. Med.* **88**, 69–84 (1966).
- Yang, L. & Botchan, M. J. *Virology* **64**, 5903–5911 (1990).
- Syväoja, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 6664–6668 (1990).
- Stillman, B. & Gluzman, Y. *Molec. cell. Biol.* **5**, 2051–2060 (1985).
- Wobbe, C. R., Dean, F. B., Murakami, Y., Weissbach, L. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4612–4646 (1986).
- Knight, J. Li, R. & Botchan, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3204–3208 (1991).
- Spalholz, B., Yang, Y. & Howley, P. M. *Cell* **42**, 183–191 (1985).

- Fuller, R., Funnell, B. & Kornberg, A. *Cell* **38**, 889–900 (1984).
- Dodson, M., McMacken, R. & Echols, H. *J. biol. Chem.* **264**, 10719–10725 (1989).
- Mastrangelo, I. *et al. Nature* **338**, 658–662 (1989).
- Borowicz, J. A., Dean, F. B., Bullock, P. A. & Hurwitz, J. *Cell* **60**, 181–184 (1990).
- Parsons, R. E., *et al. J. Virol.* **65**, 2798–2806 (1991).
- Cheng, L. & Kelly, T. J. *Cell* **59**, 541–551 (1989).
- Sharp, P. A. *Nature* **351**, 16–18 (1991).
- Shope, R. E. & Hurst, E. W. *J. exp. Med.* **58**, 607 (1933).
- Zur Hausen, H. & Schneider, A. in *The Papillomaviruses* (eds Howley, P. M. & Salzman, N.) 245–265 (Plenum, New York, 1988).
- Shaw, K. & Howley, S. in *Virology* (eds Fields, B. N. & Knipe, D. M.) 1652–1678 (Raven, New York, 1990).
- Grussenmeyer, T., Scheidtmann, K., Hutchinson, M. A., Eckhart, W. & Walter, G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7952–7954 (1985).
- Li, J. J. & Kelley, T. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6973–6977 (1984).
- Lin, S. Y. & Riggs, A. D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 947–951 (1974).

ACKNOWLEDGEMENTS. We thank L. Conroy and Sun Yung Kim for technical assistance in the growth and preparation of the recombinant baculoviruses, and H. Echols and D. Carroll for suggestions. The DNA topoisomerase I inhibitor camptothecin and topoisomerase II inhibitor VM-26 used in these studies were gifts from L. F. Liu (Johns Hopkins). The research was supported by grants from the NIH and the American Cancer Society to M.B.L.Y. was a fellow of the Damon Runyon/Water Winchell Cancer Fund, and I.M. is an American Cancer Society Fellow.

Formation of giant molecular clouds and helical magnetic fields by the Parker instability

Kazunari Shibata*† & Ryoji Matsumoto‡

* Department of Physics and Astronomy, Aichi University of Education, Kariya, Aichi 448, Japan

‡ College of Arts and Sciences, Chiba University, Chiba 260, Japan

USING the Nagoya telescope¹, Uchida *et al.*² found an unusual helical filamentary structure, spinning about its long axis, in the L1641 cloud in the Orion cloud complex. Noting that this structure is consistent with a helically twisted magnetic field inferred from optical polarization observations^{3,4}, they argued that the helical filament is a manifestation of torsional magnetohydrodynamic (Alfvén) waves draining angular momentum from a nearby massive cloud, thus promoting collapse and star formation. Here we present an alternative interpretation. We suggest that the Orion molecular cloud complex formed through the Parker instability⁵ (the buoyancy of a magnetic field entrained in matter), and that the helical filament is the result of spinning gas falling along the magnetic field and twisting it. The twisted magnetic field, unlike a purely planar field, suppresses the Parker instability on small scales, allowing the generation of finite clouds rather than general turbulence.

The Parker instability⁵ is a magnetohydrodynamic instability which occurs if a gas layer in a gravitational field is supported in part by horizontal magnetic fields. Suppose that magnetic field lines are disturbed and begin to undulate. The mass in the raised portion of a loop drains down along the field lines so that the loop top becomes lighter than the ambient medium. If the buoyancy at the loop top is larger than the restoring magnetic tension, the loop rises further and the instability sets in. Parker⁵ noted that the falling mass accumulates in the magnetic valleys, explaining the formation of interstellar cloud complexes. According to linear analysis⁵, the most unstable perturbation wavelength for $\beta = p_{\text{gas}}/p_{\text{mag}} \approx 1$, where p is pressure, is $\lambda_{\text{max}} \approx 14H \approx 1.4(H/100 \text{ pc}) \text{ kpc}$, with a growth time of $\tau_{\text{Parker}} \approx (2-5) \times (H/V_A) \approx (2-5) \times 10^7 (H/100 \text{ pc}) (V_A/10 \text{ km s}^{-1})^{-1}$ years. Here, H is the scale height of the galactic disk, $V_A = B/(4\pi\rho)^{1/2} \approx 10(B/3 \mu\text{G}) (n/1 \text{ cm}^{-3})^{-1/2} \text{ km s}^{-1}$ is the Alfvén speed, and n is the number density of atomic hydrogen. The mass, M , of the condensed gas clouds is $M \approx 7 \times 10^5 (H/100 \text{ pc})^3 (n/1 \text{ cm}^{-3}) M_{\odot}$.

Since the pioneering work by Parker, workers have continued to develop the linear theory of the Parker instability⁶⁻⁸, and its application to the formation of interstellar cloud complexes^{9,10}. The pure Parker instability (without self-gravity) has, however, been criticized as leading to chaotic structures and turbulence rather than coherent cloud complexes^{7,11}. This is because in three dimensions the most unstable mode is that with $k_{\perp} = \infty$, where $k_{\perp}$ is the wavenumber perpendicular to the magnetic fields⁵. Elmegreen⁷ pointed out that this difficulty is resolved by considering the effect of the self-gravity (the Parker-Jeans instability). We suggest another way to resolve this difficulty: if the magnetic field is sheared or twisted, the short-wavelength 'flute' mode ($k_{\perp} = \infty$) is stabilized or at least its growth rate is greatly reduced, so that a large-scale flux tube structure with finite $k_{\perp}$ is created. It is known that dynamo action usually produces a twisted field¹². Here we give an example of another mechanism for creating a twisted magnetic field.

Matsumoto *et al.*¹³ performed a self-consistent, two-dimensional, nonlinear numerical simulation of the Parker instability and found that (1) the maximum velocity of the flow down

the loop is comparable to the initial Alfvén speed, (2) shock waves are generated at the foot of the loop if $\beta(t=0) < 3$, whereas a nonlinear oscillation is excited if $\beta(t=0) > 3$, and (3) when $\beta < 3$, a quasi-static dense cloud, corresponding to a giant molecular cloud complex, is formed just below the shock front (Fig. 1). More recently, Matsumoto *et al.*¹⁴ showed that even when $\beta > 3$, shock waves are generated as a result of the nonlinear mode coupling during the nonlinear magnetic loop oscillations, and that a quasi-static cloud is eventually created. Hence, these nonlinear results¹³⁻¹⁵ suggest that if giant molecular cloud complexes are created by the Parker instability, there should be a pair of shock fronts just above (or below) the cloud complexes (Fig. 1). In the regions just behind the shocks there may be efficient formation of molecular clouds and stars, similar to those behind the galactic spiral shock waves¹⁶.

Figure 2 shows a map of the molecular clouds in Orion¹⁷. We note that the distribution pattern of the molecular cloud is similar to the V pattern of the shock fronts found in the nonlinear simulations. The size of the cloud complex (~ 100 – 200 pc) roughly agrees with that found in the simulations. Moreover, the magnetic field configuration inferred from the polarization of the star light^{3,18} is roughly consistent with that predicted by the simulation. Consequently, we suggest that the Orion molecular cloud complex might have been created by the Parker instability.

There are, however, problems with this hypothesis. First, according to the Parker instability hypothesis, the mass density is maximum inside the V pattern (Fig. 1). However, there is not much mass (in the form of molecular clouds) inside the Orion V pattern. Second, Uchida *et al.*² found a prominent helical filamentary structure suggesting the presence of helically twisted

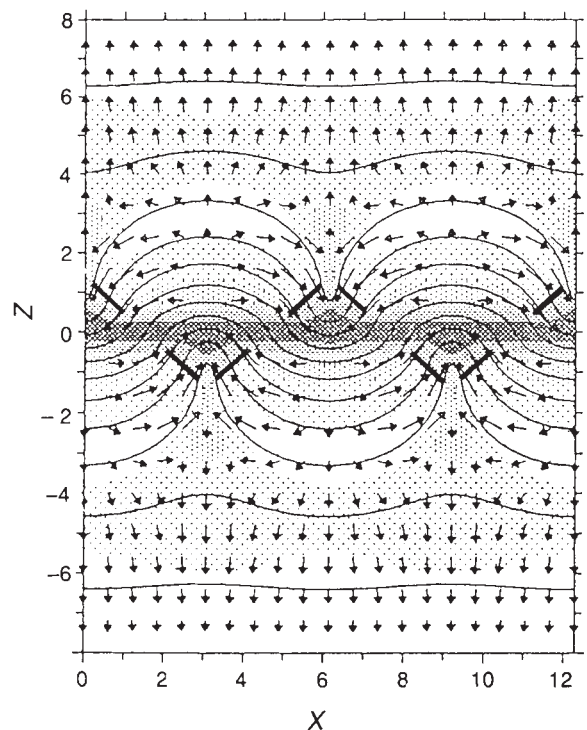


FIG. 1 A typical result of nonlinear simulations of the Parker instability¹³. In this case $\beta = p_{\text{gas}}/p_{\text{mag}} = 1$ in a later phase ($t/\tau = 8.55$). The magnetic field lines and velocity vectors are shown, and the density is indicated on a logarithmic scale (by the shaded dot pattern). The scale of the velocity vectors is linearly taken such that the maximum velocity in this figure is $2C_s$. The units of length and time are $2.3H$ and $\tau = 2.3H/C_s$, where H is the scale height of the disk and C_s is the isothermal sound speed. The maximum velocity is comparable to the initial Alfvén speed, which is slightly larger than the sound speed. The plane $z=0$ corresponds to the galactic plane. The thick solid lines, forming V or inverted V patterns, show the positions of the shock fronts produced by the supersonic downflows.

† Present address: Solar Physics Division, National Astronomical Observatory, Mitaka, Tokyo 181, Japan.

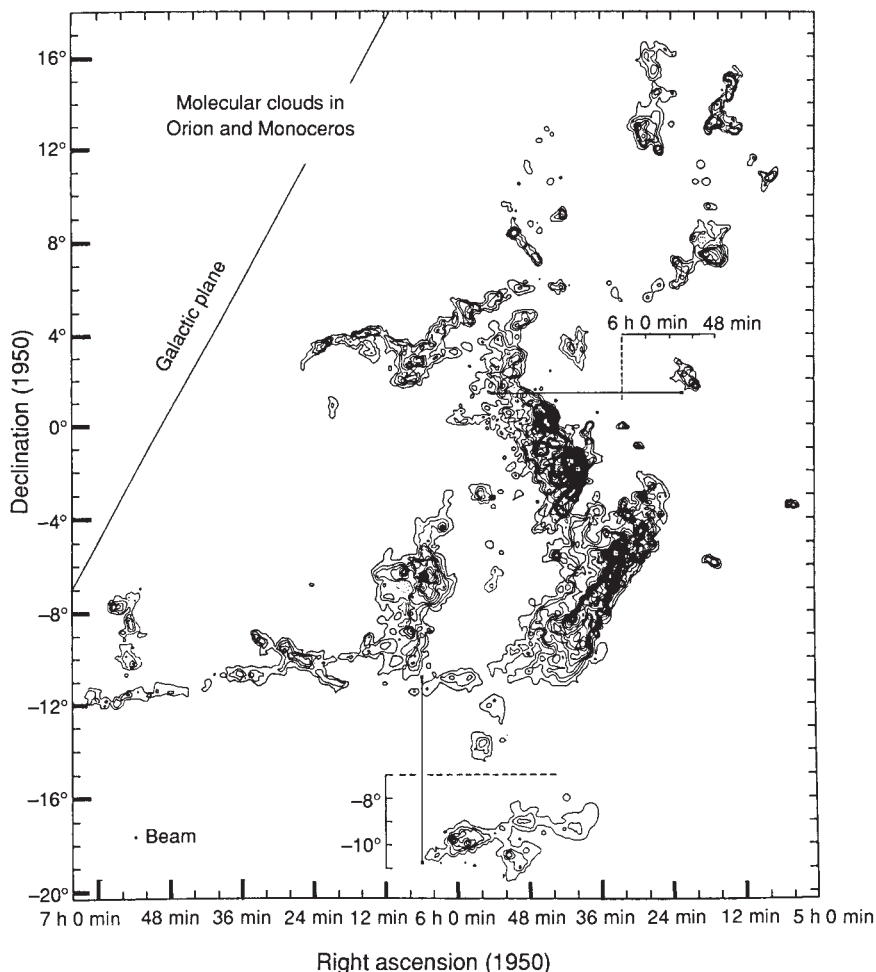


FIG. 2 Contour map of integrated CO emission in the velocity range of -10 to 20 km s^{-1} in Orion and Monoceros, taken from Maddalena *et al.*¹⁷.

magnetic field lines in the Orion molecular cloud complex. Is this newly observed structure consistent with the hypothesis of a Parker instability?

The answer to the second question is yes, as we shall discuss here. We believe (Fig. 3) that magnetic fields in giant molecular clouds can be expected to be helically twisted. These twists exert magnetic pressure, $\nabla(B^2/8\pi)$, along the magnetic flux tube¹⁹, supporting the heavy clouds. Hence it is not surprising that there is less mass than expected inside the V pattern; its fall is broken by the magnetic pressure of the twisted field. This may be an answer to the first question. (Note that the common explanation of the cloud distribution pattern in Orion is that the large pressure exerted by the old OB star association on the right pushed the molecular cloud to the left²⁰. Our new model may not be inconsistent with this hypothesis, because the V pattern itself could have been primordially produced by our mechanism.)

Nonlinear simulations¹³ show that the infalling masses are compressed both by the shock waves and by geometrical effects such as the decreasing cross-sectional area of the flux tube; in fact, the density increases by a factor of 3–10 during the nonlinear stage of the Parker instability. If we consider the effects of additional cooling mechanisms or self-gravity, the infalling masses could be further compressed and finally evolve into a molecular cloud whose density is more than 10^3 times that of the initial gas. During this contraction of the gas, the Coriolis force due to galactic rotation causes the cloud to spin (Fig. 4). This is why magnetic field lines are helically twisted as a result of the Parker instability. Because the sense of the rotation due to the Coriolis force is the same for all contracting clumps of gas in the galactic disk (Fig. 4), the magnetic twists caused by the gas clumps accumulate in the valleys, and also in the upper loops. Note that the whole magnetic valley itself is twisted by

the Coriolis force because the average density also grows in the valley. This global twist, however, is not shown in Fig. 4. Consequently we expect that there should be helically twisted flux tubes in all molecular cloud complexes produced by the Parker instability.

To estimate the total angular momentum in the infalling cloud, the angular velocities and other physical quantities, we assume that the infalling cloud is initially described by the following parameters: density $\rho_1 \approx 10^{-24} \text{ g cm}^{-3}$, size $R_1 \approx 100 \text{ pc}$ and angular velocity $\Omega_1 \approx 10^{-15} \text{ s}^{-1} \approx \Omega_{\text{Galaxy}} (r \approx 10 \text{ kpc})$. Suppose the cloud evolves into a state with density ρ_2 , size R_2 and angular velocity Ω_2 , as it falls through the shocks and into the magnetic

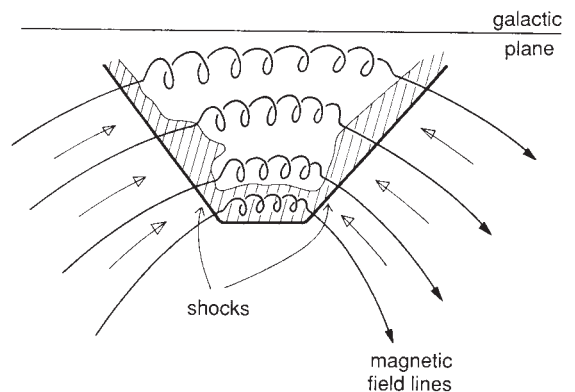


FIG. 3 Schematic diagram of the cloud complex produced by nonlinear evolution of the Parker instability. The helically twisted magnetic flux tubes are generated because of the Coriolis force, and these accumulate inside the giant molecular clouds. This picture may be compared with the observed CO map of the Orion molecular cloud complex (Fig. 2).

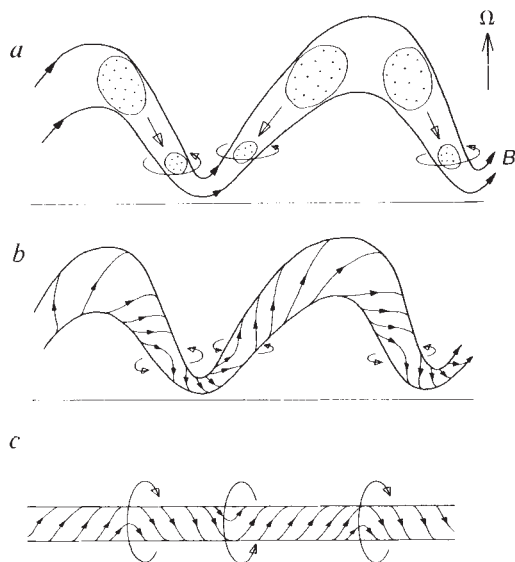


FIG. 4 Schematic diagram showing the origin of the helically twisted magnetic fields. *a*, Before the flux tube is twisted; *b*, after the tube is twisted. Note that the magnetic twist accumulates both in the magnetic field valleys and in the loop tops. *c*, Schematic illustration showing how the tube is twisted. Note that this tube is topologically the same as that shown in *b*.

field valleys. Then the total mass of the infalling cloud is $M_{\text{infall}} \approx 10^4 M_{\odot}$, which is only a fraction of the total mass involved in the cloud complex ($\sim 10^6 M_{\odot}$). The total angular momentum is $L_{\text{infall}} \approx M_{\text{infall}} R_1^2 \Omega_1 \approx 10^{63.5} \text{ g cm}^2 \text{ s}^{-1}$. Note that this is much larger than the total angular momentum observed in the helical filament and its base cloud¹ (Y. Uchida, K.S. and R. Rosner, manuscript in preparation), which is $\sim 10^{59} - 10^{60} \text{ g cm}^2 \text{ s}^{-1}$. The mass conservation law, $R_2^3 \rho_2 \approx R_1^3 \rho_1$, and angular momentum conservation law, $R_2^2 \Omega_2 \approx R_1^2 \Omega_1$, yield $\Omega_2 \approx (R_1/R_2)^2 \Omega_1 \approx (\rho_2/\rho_1)^{2/3} \Omega_1 \approx 10^{-13} \text{ s}^{-1}$, if $\rho_2 \approx 10^3 \rho_1 \approx 10^{-21} \text{ g cm}^{-3}$. This is much larger than the usual angular velocities of molecular clouds²¹, $10^{-15} - 10^{-14} \text{ s}^{-1}$. Similarly, the rotational velocity of the cloud becomes $V_{\phi,2} \approx R_2 \Omega_2 \approx (\rho_2/\rho_1)^{1/3} R_1 \Omega_1 \approx 30 \text{ km s}^{-1}$, which is much larger than the observed rotational velocities of molecular clouds, $V_{\phi} \approx 0.1 - 1 \text{ km s}^{-1}$. Consequently, we have to consider the effect of magnetic braking caused by generation of the helical magnetic twists (Fig. 4) to explain the small rotational velocities actually observed.

We shall now estimate the braking time and the degree of magnetic twist. If the Alfvén transit time across a tenuous loop is much shorter than the growth time of the Parker instability, then the braking time can be approximated by $\tau_{\text{braking}} \leq \lambda/2 \langle V_A \rangle \approx 7 \times 10^7 \text{ yr} \approx \text{a factor} \times \tau_{\text{Parker}}$, where $\langle V_A \rangle$ is the average Alfvén speed inside the whole flux tube including the loop and the valley. The ratio of the twisted component, B_{ϕ} , to the initial field strength, B_0 , becomes

$$\frac{B_{\phi}}{B_0} \approx \frac{V_{\phi,2}}{\langle V_A \rangle} \approx 3 \left(\frac{\rho_2/\rho_1}{10^3} \right)^{1/3} \left(\frac{R_1}{100 \text{ pc}} \right) \left(\frac{\Omega_1}{10^{-15} \text{ s}^{-1}} \right) \left(\frac{\langle V_A \rangle}{10 \text{ km s}^{-1}} \right)^{-1}$$

Hence the degree of the twist, B_{ϕ}/B_0 , is of the order of unity, which is comparable to the observed degree of the twist in the helical filaments in the Orion molecular cloud complex².

Finally, we note that Beck *et al.*²² found a three-dimensional, arc-like magnetic field structure extending out of M31 on a scale of $\sim 4.5 \text{ kpc}$, which also fits our model well. This field structure is associated with a large H I cloud complex and with a CO cloud²³ with a strong nonthermal radio source. Hence they suggested that this might be the first Parker-Jeans instability observed in a galaxy other than our own. Another interesting fact about this M31 cloud complex is that the CO intensity is low in its centre (that is, a double-peaked CO feature)²³, similar to the distribution pattern of the molecular clouds in the Orion complex. We suggest that the low CO intensity may correspond

to a region with strong, helically twisted magnetic fields (Fig. 3). *Note added in proof:* R. Beck has informed us that the distance between the two peaks in the M31 CO cloud²³ is about 250 pc, which is much smaller than the arc ($\sim 4.5 \text{ kpc}$) and is comparable to the distance between the two shocks in our model (Fig. 3). $\square$

Received 21 April; accepted 19 August 1991.

1. Fukui, Y., Iwata, T., Mizuno, A., Bally, J. & Lane, A. in *Protostars and Planets III* (University of Arizona Press, in the press).
2. Uchida, Y. *et al. Nature* **349**, 140–142 (1991).
3. Heiles, C. in *Interstellar Processes* (eds Hollenbach, D. J. & Thronson, H. A.) 171–194 (Reidel, Dordrecht, 1987).
4. Vrba, F., Strom, S. E. & Strom, K. *Astr. J.* **96**, 680–694 (1988).
5. Parker, E. N. *Astrophys. J.* **145**, 811–833 (1966).
6. Shu, F. H. *Astr. Astrophys.* **33**, 55–72 (1974).
7. Elmegreen, B. G. *Astrophys. J.* **253**, 634–654 (1982).
8. Nakamura, H., Hanawa, T. & Nakano, T. *Publs. Astr. Soc. Japan* (in the press).
9. Mouschovias, T. Ch., Shu, F. H. & Woodward, P. R. *Astr. Astrophys.* **33**, 73–77 (1974).
10. Blitz, L. & Shu, F. H. *Astrophys. J.* **238**, 148–157 (1980).
11. Asseo, E., Cesarsky, C. J., Lacheze-Rey, M. & Pellat, R. *Astrophys. J.* **225**, L21–L25 (1978).
12. Parker, E. N. *Cosmical Magnetic Fields*, 532–611 (Clarendon, Oxford, 1979).
13. Matsumoto, R., Horiuchi, T., Shibata, K. & Hanawa, T. *Publs. Astr. Soc. Japan* **40**, 171–195 (1988).
14. Matsumoto, R., Horiuchi, T., Hanawa, T. & Shibata, K. *Astrophys. J.* **356**, 259–271 (1990).
15. Shibata, K. *et al. Astrophys. J.* **338**, 471–492 (1989).
16. Fujimoto, M. in *IAU Symp.* No. 29 (ed. Arakeljan, M.) 453–463 (Izdatel'stvo Akademii Nauk Armyanski SSR, 1968).
17. Maddalena, R. J., Morris, M., Moscowitz, J. & Thaddeus, P. *Astrophys. J.* **303**, 375–391 (1986).
18. Appenzeller, I. *Astr. Astrophys.* **36**, 99–105 (1974).
19. Shibata, K. & Uchida, Y. *Publs. Astr. Soc. Japan* **37**, 31–46 (1985).
20. Bally, J., Langer, W. D., Stark, A. A. & Wilson, R. W. *Astrophys. J.* **312**, L45–L49 (1987).
21. Goldsmith, P. F. & Arquilla, R. in *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) 137–149 (University of Arizona Press, Tucson, 1985).
22. Beck, R., Loiseau, N., Hummel, E., Berkhuijsen, E. M. & Wielebinski, R. *Astr. Astrophys.* **222**, 58–68 (1989).
23. Bajaja, E., Berkhuijsen, E. M. & Beck, R. in *IAU Symp.* No. 140 (eds Beck, R. *et al.*) 203–204 (Kluwer, Dordrecht, 1990).

ACKNOWLEDGEMENTS. We thank the referee, B. Elmegreen, for helpful comments, A. McAllister for his help in revising the manuscript, and R. Beck, M. Chiba, M. Fujimoto, Y. Fukui, T. Hanawa, T. Hasegawa, N. Nakai, S. Miyama, T. Nakano, T. Sawa, Y. Sofue, K. Tatematsu and Y. Uchida for discussions.

Detection of a stellar flare at extreme ultraviolet wavelengths

M. A. Barstow*, G. E. Bromage†, G. S. Pankiewicz*§, R. González-Riestra‡, M. Denby* & J. P. Pye*

* Department of Physics and Astronomy, University of Leicester

† Space Astronomy Group, Rutherford Appleton Laboratory, Didcot, Oxfordshire, UK

‡ IUE Observatory, Vilspa, Madrid, Spain

THE transition region between stellar chromospheres and coronae contains plasma at temperatures of $10^5 - 10^6 \text{ K}$, radiating predominantly at extreme ultraviolet (EUV) wavelengths (60–1,000 Å; 0.012–0.2 keV). To understand the energy transport processes that maintain the corona, radiative losses from this region must be studied, particularly during the dramatic energy release that occurs in a stellar flare. During the all-sky survey conducted by the Rosat Wide Field Camera¹, we monitored the binary flare star system BY Draconis, with coverage by the International Ultraviolet Explorer satellite far ultraviolet and optical observations and by the Rosat X-ray telescope² for part of the time. We detected a stellar flare in all four wavebands. This is the first unambiguous extreme ultraviolet detection of a flare, and one of the widest simultaneous wavelength-range coverages obtained. The peak luminosity and total energy of this flare, in the photon energy range 0.08–0.18 keV, are comparable with the values obtained for a number of flares integrated over a larger energy range (0.05–2.0 keV) by Exosat satellite observations in 1983–86. We conclude that radiation in the extreme ultraviolet carries away a substantial fraction of the total flare energy.

The launch of the Wide Field Camera (WFC) telescope¹ in June 1990, on board the Rosat² mission, has allowed the first

§ Present address: Wide Field Camera Quick Look Facility, MPE, Garching, Munich, Germany.

all-sky survey to be performed in the extreme ultraviolet (EUV) waveband. With the large 5° field of view of the WFC, the survey strategy not only enabled large numbers of new EUV sources to be detected but also allowed monitoring of their luminosity for several days. The nature of the survey allowed the prediction of viewing 'windows' for interesting objects well in advance of the observations, presenting an opportunity of coordinating Rosat coverage with other ground- and space-based observatories. A large block of International Ultraviolet Explorer (IUE) satellite time was set aside for this purpose under the 'RIASS' programme. We present here Rosat WFC and IUE observations of the prototype 'spotty' binary flare star BY Draconis, during which it was observed to flare by both instruments. The flare was also detected by the Rosat X-ray telescope (J.H.M.M. Schmitt and T.A. Fleming, personal communication).

These data represent an unambiguous detection of an EUV flare in a star other than the Sun and also the first observation of any flare simultaneously in optical, ultraviolet, EUV and X-ray bands. We concentrate here on a detailed study of the EUV flare, determining the energy release in this band and the importance of this as an aid to understanding the general nature of stellar flares.

BY Draconis is a double lined spectroscopic binary with a period of 6 days, comprising stars of spectral type dM0e and dM1-2e. Its optical photometric period of 3.8 days is assumed to be the result of light modulation by large spot areas on the more active primary, although flares have been observed on both components. It is clear that the stars' rotations are not tidally locked to the orbital period, and the system seems to be relatively young, still approaching the main sequence³. The radius of the primary is slightly larger than main sequence at $0.9 M_\odot$ (compare this with 0.6 for an M0 dwarf), and the mass ratio is 1.02.

BY Draconis was observed by the WFC for a period of 19 days from 23 September to 11 October 1990. In the WFC, the EUV wavebands are determined by thin-film filters. Four different filters are available, but only two were used during the survey, denoted S1 and S2, covering the wavelength ranges

60–140 Å and 114–180 Å respectively. The filters were exchanged once per day, and the BY Dra observations were therefore conducted first in one band and then the other.

Six IUE shifts from the RIASS programme were allocated to BY Dra, and these were timed to coincide with some of the optimum WFC exposures, from 29 September to 4 October, giving 7 hours of ultraviolet observations during each day. To optimize the use of the available time, alternate spectra were taken with the short-wavelength (1,150–1,950 Å, SWP) and long-wavelength (1,900–3,250 Å, LWP) cameras. In this approach, one camera is being read and prepared for its next exposure while the other is being exposed. Between the far-ultraviolet spectra the system's optical brightness was measured with the IUE fine error sensor (FES), which yields a magnitude estimate between V and B bands. In this paper we present measurements of the flux in the 1,550 Å line of CIV, the clearest indicator of stellar transition region activity, together with the FES measurements, for comparison with the WFC data.

The IUE data were processed using the standard, well-established data analysis packages IUEDR⁴ and DIPSO⁵. The unusual nature of the WFC data, however, requires a brief description. As a result of the scan strategy used during the sky survey, the data arrived as a series of discrete ~50-s 'snapshots' of the source separated in time by one ROSAT orbit (5,760 s). WFC survey data are archived in files containing all the event timing and position information, from which the photon events can be recovered and a time series produced. First, a time series of the events in a circle of radius 6 arcmin centred on the source position (encompassing the point response of the instrument) is created separately for each filter. Second, the background contribution is estimated for each filter from data lying in an annulus surrounding the source. The raw time series are then corrected for deadtime, exposure and telescope vignetting.

Figure 1a shows the background-subtracted EUV light curve of BY Dra extracted from the WFC survey data. To optimize the signal-to-noise in the data, each time bin is an average of three of the once-per-orbit 'snapshots'. The mean source count rates are 0.051 ± 0.004 counts s^{-1} and 0.057 ± 0.004 counts s^{-1} in

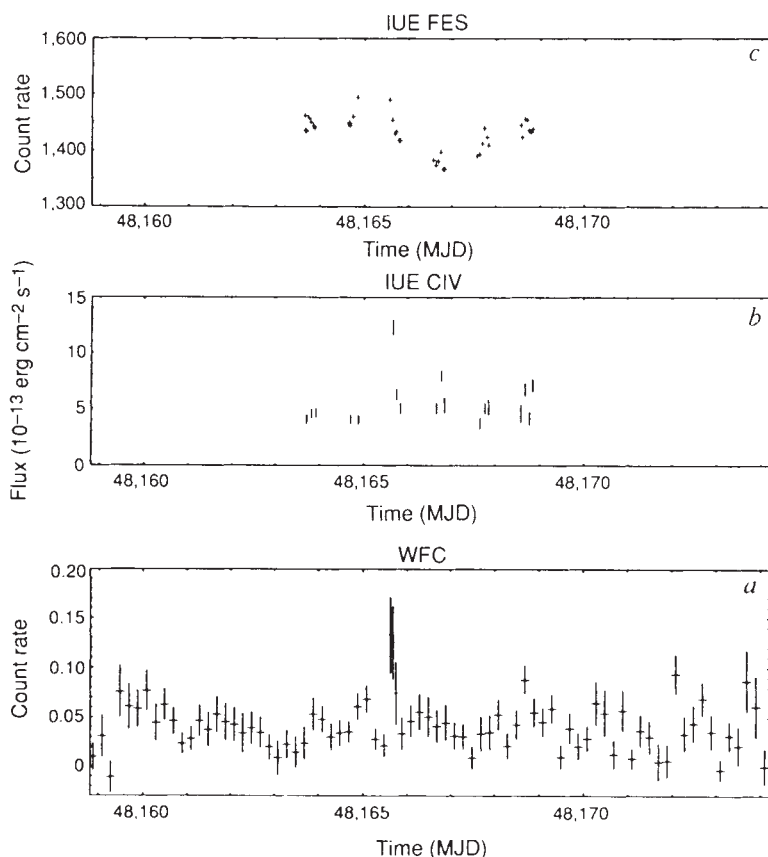


FIG. 1 Fifteen days of EUV observations of BY Draconis (a) recorded during the Rosat all-sky survey. Simultaneous far-ultraviolet (b) and optical (c) data from IUE span the central six days of the survey coverage.

S1 and S2 filters respectively. Because these rates are identical within the experimental error, we have been able to merge the separate S1 and S2 time series to produce a single, composite light curve here. Two distinct features can be seen in the EUV light curve. First, there is a clear and possibly periodic modulation of the EUV flux. Second, there is a strong flare, occurring at a minimum in the quiescent flux, between MJD (modified Julian dates) 48165 and 48166. The peak count rate in the flare is a factor of 3.5 higher than the quiescent level beforehand, allowing this section of the data to be displayed with its full, single-orbit, resolution. The EUV modulation is likely to be related to the photometric wave or rotation of the system, but this will not be considered here.

The flare shown in Fig. 1a is the first to be detected in the EUV on any star other than the Sun. The flare can also be seen in the far-ultraviolet and optical band. The integrated flux in the CIV line (Fig. 1b) shows an increase of a factor of 2.5, and the FES count (Fig. 1c) indicates a change of ~ 0.04 mag in a wavelength band between B and V, probably corresponding to a U-band enhancement of ~ 0.5 –1.0 mag.

We have folded a set of single-temperature models for the emission from an optically thin plasma (ref. 6) through the instrument response of each WFC filter. Normalizing these to the observed count rates in each filter and taking into account the attenuation of the interstellar medium allows the EUV flux to be determined as a function of the model temperature. For a given value of the interstellar column density (N_H), the point at which the predicted S1 and S2 fluxes are identical corresponds to a model temperature that is consistent with the data. At a distance of 15.63 pc, BY Dra lies within the local bubble of the interstellar medium. Estimates of N_H from the average local volume density (0.07 cm^{-3}) and also contour maps of the column density^{7–9} give a value of $3 \times 10^{18} \text{ cm}^{-2}$. Adopting this value in our analysis gives us a temperature of $\sim 3 \times 10^6 \text{ K}$ for the plasma responsible for the quiescent EUV emission in this system. The corresponding luminosity and emission measure are $3.9 \times 10^{28} \text{ erg s}^{-1}$ and $2.0 \times 10^{52} \text{ cm}^{-3}$ in the energy range 0.08–0.18 keV. Uncertainties in these derived values are dominated by the limitations in the accuracy of the spectral models in predicting fluxes and are about $\pm 50\%$ (ref. 10).

It is not possible to repeat the above analysis to determine the temperature of the flare plasma, as the event was only observed in the S2 filter. But by assuming that the plasma temperature remains constant, it is possible to make first-order estimates of the parameters of the flare. The peak count rate of $0.11 \text{ counts s}^{-1}$ corresponds to an emission measure of $4.0 \times 10^{52} \text{ cm}^{-3}$, twice the mean value. If the flare cools radiatively, its duration, lasting roughly three Rosat orbits in the EUV ($17,000 \pm 3,000 \text{ s}$; see Fig. 1a), allows the electron density (n_e) in the plasma to be estimated. Taking $T = 3 \times 10^6 \text{ K}$ gives a density of 10^{10} cm^{-3} , with an error of $\pm 30\%$ from the combined uncertainty in the temperature determination and flare duration. Finally, the total energy radiated in the EUV (0.08–0.18 keV) was $6.8 \times 10^{32} \text{ erg}$, integrated over the duration of the flare.

The simultaneous optical and far-ultraviolet data demonstrate unambiguously that the WFC has detected a flare in the EUV. EUV emission has only once before been reported from a dMe flare star¹¹, Proxima Centauri. Because in that study the star was only detected on one of two brief observations, it was assumed that a flare had occurred, but there was no other evidence to support this assertion. Furthermore, for the Prox Cen observation only a luminosity integrated across the wide band 0.05–1 keV is reported. Thus it is not possible to compare the Prox Cen and BY Dra observations directly. Fortunately, dMe flares have been observed at X-ray energies in the past, and we can refer to these in discussing the importance of our observation. Pallavicini *et al.*¹² have recently reviewed many Exosat observations of flare stars. Although the Exosat low-energy experiment extended into the EUV (down to 0.05 keV), the observed count rates were dominated by the soft X-ray

spectrum, because of the relatively high energy (2.0 keV) mirror cutoff, and so any purely EUV component could not be separately estimated from the Exosat data. X-ray flares have been observed with peak luminosities (0.05–2 keV) in the range 10^{28} – $5 \times 10^{30} \text{ erg s}^{-1}$ with total energies in the range 3×10^{30} – 10^{34} erg .

Flares can be phenomenologically divided into two groups, impulsive flares where the decay times are tens of minutes and long decay flares with decay times of an hour or longer, reminiscent of two-ribbon solar flares. The BY Dra flare observed by the WFC seems to fall into the latter category, and it is interesting that the decay time of $\sim 4.5 \text{ h}$ is longer than any of those observed by Exosat. The peak luminosity and total energy (in the range 0.08–0.18 keV) of the EUV flare are comparable with many of those observed over a much larger energy range (0.05–2 keV) by Exosat. The single flare observed on BY Dra by Exosat had a total energy of $1.2 \times 10^{33} \text{ erg}$, only a factor of two greater than that of the BY Dra flare reported here for the much narrower energy range purely in the EUV.

We conclude that radiation in the EUV carries away a substantial fraction of the total energy released in such a flare. Therefore, EUV observations are important if the mechanisms of energy release in stellar flares are to be fully understood. Further studies made with the Rosat WFC, from analysis of the existing sky-survey data base and the subsequent pointed observations of specific targets, will make a considerable contribution to this field. $\square$

Received 8 July; accepted 30 August 1991.

1. Wells, A. *et al. Proc. SPIE* **1344**, 230–243 (1990).
2. Pfefferman, E. *et al. Proc. SPIE* **733**, 519–532 (1986).
3. Vogt, S. S. & Fekel, F. *Astrophys. J.* **234**, 958–963 (1979).
4. Giddings, J. & Rees, P. *Starlink User Note* **37** (1989).
5. Howarth, I. D. & Murray, J. *Starlink User Note* **50** (1990).
6. Raymond, J. C. & Smith, B. W. *Astrophys. J. Suppl.* **35**, 419–440 (1977).
7. Frisch, P. C. & York, D. G. in *Extreme Ultraviolet Astronomy* (eds Malina, R. F. & Bowyer, S.) 322–332 (Pergamon, New York, 1989).
8. Frisch, P. C. & York, D. G. *Astrophys. J.* **271**, L59–64 (1983).
9. Paresce, F. *Astr. J.* **89**, 1022–1037 (1984).
10. Raymond, J. C. in *Hot Thin Plasmas in Astrophysics* (ed. Pallavicini, R.) 3–20 (Kluwer, Dordrecht, 1988).
11. Haisch, B. M., Linsky, J. L., Lampton, M., Paresce, F., Margon, B. & Stern, R. *Astrophys. J.* **213**, 119–124 (1977).
12. Pallavicini, R., Tagliaferri, G. & Stella, L. *Astr. Astrophys.* (in the press).

ACKNOWLEDGEMENTS. M.A.B., G.E.B., J.P.P., G.S.P. and M.D. acknowledge the support of the SERC through its funding of the Rosat WFC project and the NASA-ESA-SERC IUE project. M.A.B. also acknowledges the support of a SERC Advanced Fellowship. The Rosat and IUE data reduction were performed on the Leicester and RAL Nodes of the SERC Starlink computer network, using Starlink hardware and software. R.G. is supported by the ESA IUE project.

Phosphates in pallasite meteorites as probes of mantle processes in small planetary bodies

Andrew M. Davis* & Edward J. Olsen†

* Enrico Fermi Institute, † Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

THE stony-iron pallasite meteorites, which are generally believed to come from the core–mantle boundaries of asteroids, contain small quantities of phosphate minerals. Here we use trace element analyses of these phosphates to investigate the magmatic history of the silicate portions of pallasites. In Eagle Station and seven other pallasites, the phosphates have relatively low concentrations of rare earth elements (REE) and are strongly enriched in heavy relative to light REE. These patterns are consistent with formation of phosphate by subsolidus reactions between metal and silicate, in which phosphate inherits the REE pattern of olivine. In Springwater and Santa Rosalia, calcium-rich phosphates have higher concentrations of REE, are enriched in light relative to heavy REE and have negative europium anomalies. These patterns are

TABLE 1 Chemical composition of olivine and phosphates in Eagle Station and Springwater

	1	2	3	4	5	6	7	8
Na ₂ O	0.001	0.000	1.55	0.000	0.000	0.000	2.50	3.56
MgO	42.32	20.88	3.59	44.05	42.84	20.86	3.61	31.26
Al ₂ O ₃	0.002	0.002	0.002	0.006	0.003	0.002	0.003	0.002
SiO ₂	38.91	0.026	0.022	38.88	0.022	0.030	0.026	2.33
P ₂ O ₅	0.005	50.28	47.27	0.007	53.46	49.34	46.92	48.92
K ₂ O	0.002	0.003	0.002	0.001	0.002	0.001	0.070	0.017
CaO	0.018	25.61	48.15	0.005	0.077	25.58	47.30	10.25
Cr ₂ O ₃	0.030	0.006	0.006	0.024	0.012	0.023	0.011	0.015
MnO	0.179	0.301	0.025	0.329	0.225	0.488	0.027	0.850
FeO	18.86	4.73	0.608	17.13	4.09	4.43	0.913	4.95
Sum	100.33	101.83	101.23	100.44	100.73	100.75	101.37	102.15
Ion microprobe spots								
	4	1	6	2	4	5	3	4
Y	<0.012	0.121 ± 0.035	2.79	0.0170 ± 0.0085	0.162 ± 0.018	17.7	206	67.9
La	<0.0040	0.0046 ± 0.0044	0.0320 ± 0.0044	0.0052 ± 0.0029	<0.0030	17.7	64.8	6.16
Ce	<0.0072	<0.018	0.095 ± 0.012	<0.0077	<0.0067	32.3	145	14.4
Pr	<0.012	<0.022	<0.012	<0.013	<0.0096	4.94	27.6	2.46
Nd	<0.042	<0.080	0.170 ± 0.035	<0.048	<0.035	24.9	165	14.6
Sm	<0.018	<0.026	0.122 ± 0.014	<0.014	<0.014	6.21	51.7	5.23
Eu	<0.0034	<0.0089	0.0840	<0.0040	<0.0032	0.492	2.40	0.306
Gd	<0.021	<0.058	0.236 ± 0.029	<0.026	<0.019	6.04	42.7	8.37
Tb	<0.0058	<0.0094	0.0355 ± 0.0067	<0.0056	<0.0052	0.794	5.80 ± 0.63	1.66
Dy	<0.020	<0.035	0.393	<0.026	0.0139 ± 0.0086	4.49	43.6	11.4
Ho	<0.010	<0.014	0.088 ± 0.012	<0.0085	<0.0070	0.809	8.27	2.43
Er	<0.013	<0.030	0.385	<0.014	0.0270 ± 0.0081	1.68	19.3	7.59
Tm	<0.0038	<0.0088	0.118	<0.0053	0.0070 ± 0.0026	0.263	3.01	1.29
Yb	<0.015	0.065 ± 0.029	1.15	<0.020	0.0355 ± 0.0096	1.07 ± 0.14	18.5	8.78
Lu	<0.0050	0.0197 ± 0.0086	0.199	<0.0059	0.0073 ± 0.0028	0.152 ± 0.037	2.98 ± 0.41	1.63

Oxide concentrations, determined by electron microprobe, are given in wt%. Elemental concentrations, determined by ion microprobe, are given in p.p.m. The uncertainties in ion microprobe analyses are $\pm 1\sigma$ based on counting statistics and are only given when they exceed 10% of the concentration. Upper limits are $<2\sigma$. 1, Eagle Station olivine; 2, Eagle Station stanfieldite; 3, Eagle Station whitlockite; 4, Springwater olivine; 5, Springwater farringtonite; 6, Springwater stanfieldite; 7, Springwater whitlockite (ion microprobe analyses corrected for overlap onto metal and silicophosphate); 8, Springwater silicophosphate mineral.

consistent with crystallization of phosphate from a europium-depleted chondritic liquid. This is unlikely to have happened near the base of the differentiating parent-body mantle, because phosphates are late-crystallizing phases; this suggests that some pallasites may come from regions of their parent bodies much nearer the surface than the core-mantle boundary.

The pallasites are stony-iron meteorites consisting of iron-nickel and olivine with minor troilite, schreibersite, chromite, phosphates and pyroxenes. They are believed to be samples of the core-mantle boundaries of small parent bodies, in which cumulate olivine and chromite mixed with molten metal^{1,2}. The pallasites have been classified on the basis of chemical composition of metal. Most belong to the 'main group', and their metal contains 8–11.5 wt% Ni and 0.01–5 p.p.m. Ir (refs 1, 2). Eagle Station and Itzawisis contain ~15 wt% Ni and 13–20 p.p.m. Ir (refs 1, 2) and have oxygen isotopic compositions similar to C3 chondrites³, whereas all other pallasites measured have oxygen isotopic compositions similar to basaltic achondrites³. Most pallasites contain one or more of the phosphate minerals stanfieldite [Mg₃Ca₃(PO₄)₄], whitlockite [Ca₃(PO₄)₂] and farringtonite [Mg₃(PO₄)₂]⁴, which may have been cumulate phases along with olivine or may have formed by oxidation of metallic phosphorus or schreibersite during or after mixing of olivine and metal⁵. The REE can be highly concentrated in phosphate minerals, whereas they are virtually excluded from the close-packed olivine crystal structure. The REE are strongly lithophile elements and are essentially completely excluded from metal under the redox conditions encountered in the Solar System. Phosphates formed in the silicate mantles of the pallasite parent bodies would be high in REE because they would concentrate REE from a source of roughly chondritic composition. Phosphates formed by redox reaction after olivine-metal mixing should be low in REE because they would inherit REE from olivine.

We have found phosphates in 10 of 17 pallasites studied so far. The concentrations of REE and other trace elements in individual phases in these 10 pallasites were determined by ion microprobe. Despite differences in oxygen isotopic and chemical composition, Eagle Station, Itzawisis and the main-group pallasites Admire, Albin, Antofagasta, Finmarken, Imilac and Mount Vernon, have qualitatively similar REE patterns in their stanfieldite and whitlockite. Eagle Station will be discussed in detail here as a representative of this group, because it contains more abundant phosphate and also shows textural evidence for the origin of the phosphate. The main-group pallasite Santa Rosalia and the unusually phosphate-rich main-group pallasite Springwater have similar REE patterns in stanfieldite. Springwater is treated in detail here because it contains a wider variety of phosphate minerals.

In the polished section of Eagle Station studied here, about half of the olivine (Fa₂₁) grains have 100–600- μ m-thick rims consisting of intergrowths of clinopyroxene (Fs₇Wo₄₅, 0.20 wt% Al₂O₃), orthopyroxene (Fs₁₇Wo₁, 0.11 wt% Al₂O₃), troilite, whitlockite and stanfieldite. Whitlockite or stanfieldite is usually in contact with metal and is separated from clinopyroxene by a 10–20- μ m-thick band of orthopyroxene. Whitlockite and stanfieldite grains can be up to 200 μ m across.

The occurrence of phosphates in Springwater is different. Buseck⁶ reported that Springwater contains ~4 vol% farringtonite and minor stanfieldite, whitlockite and phosphoran olivine. Farringtonite occurs as large areas replacing metal as the phase between rounded olivine grains (Fa₁₈). One large anhedral stanfieldite grain, 1 mm across, was found between olivine and metal in one of the four polished sections studied. A 300 × 600- μ m rounded crystal of a new magnesium-calcium-sodium silicophosphate mineral (Table 1, sample 8) was attached to an olivine grain and in contact with troilite and schreibersite. A 30 × 100 μ m grain of whitlockite was found

between the silicophosphate and schreibersite. No phosphoran olivine was found.

In the remaining pallasites in which phosphates were found, stanfieldite or whitlockite or both occur as 50- μm to 1-mm grains between olivine and metal or troilite. Most pallasites have been more strongly thermally metamorphosed than Eagle Station⁷, such that textures like that seen in Eagle Station would have been obliterated.

Major and minor elements were determined with a Cameca SX-50 electron microprobe using wavelength-dispersive spectrometers. Trace-element data were collected with an AEI IM-20 ion microprobe using magnetic peak switching at low mass resolution ($M/\Delta M = 500$). A mass-filtered 10-kV primary beam of $^{16}\text{O}^+$ ions was operated at a current of 10–30 nA and produced a spot 15–30 μm in diameter. Molecular interferences were suppressed by energy filtering. Calcium-normalized ion yields and interferences were determined from silicate minerals and glasses. Silicates and phosphates are known to have similar calcium-normalized ion yields^{8,9}.

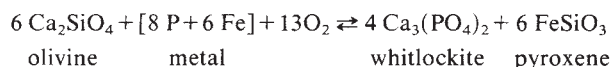
Some analyses of Eagle Station and Springwater are shown in Table 1. Each ion microprobe analysis is the average of several spots (numbers are given in Table 1). No significant variation in composition was found from spot to spot. The C1 chondrite-normalized REE patterns of all phases analysed are shown in Fig. 1. Yttrium is plotted with REE because it is similar in geochemical behaviour to dysprosium and holmium and is determined more precisely because of its higher cosmic abundance.

Phosphates in Eagle Station are remarkably low in REE. Whitlockite is enriched in heavy REE relative to light REE with a smoothly fractionated pattern and a small positive europium anomaly. Stanfieldite has a REE pattern similar in shape to that of the whitlockite, but has C1 chondrite-normalized enrichment factors that are lower by factors of 10–20. REE and yttrium

were not detected in Eagle Station olivine: 2σ upper limits for enrichment factors relative to C1 chondrites range from <0.008 to <0.2 .

Three of the phosphates found in Springwater are highly enriched in REE, whereas the fourth, farringtonite, is very low in REE and has a heavy REE-enriched pattern similar to that of Eagle Station stanfieldite. Whitlockite and stanfieldite in Springwater have light REE-enriched patterns, whereas the silicophosphate has a heavy REE-enriched pattern. Stanfieldite, whitlockite and the silicophosphate all have pronounced negative europium anomalies. REE have been measured in whitlockite in equilibrated ordinary chondrites^{10,11}, and in the Angra dos Reis and ALH 86010 angrite achondrites¹². The shape and enrichment factors of the REE pattern in Springwater whitlockite (enrichment factor for La = 276; La/Sm = 0.79 relative to C1 chondrites; Eu/Eu* = 0.15, are similar to those of whitlockite in ordinary chondrites (La = ~200; La/Sm = ~0.75; Eu/Eu* = 0.11). Angrite whitlockite has higher REE enrichment factors (La = 360–820), fractionated light REE patterns (La/Sm = ~3.5) and smaller negative europium anomalies (Eu/Eu* = ~0.7).

The textures found in the phosphate-pyroxene-troilite rims on many Eagle Station olivines suggest that the rims may have formed by subsolidus reactions between olivine and metal. Whitlockite and pyroxene could form by the following reaction:



Similar reactions can be written for the formation of stanfieldite and calcic pyroxene. Oxygen in these subsolidus reactions was

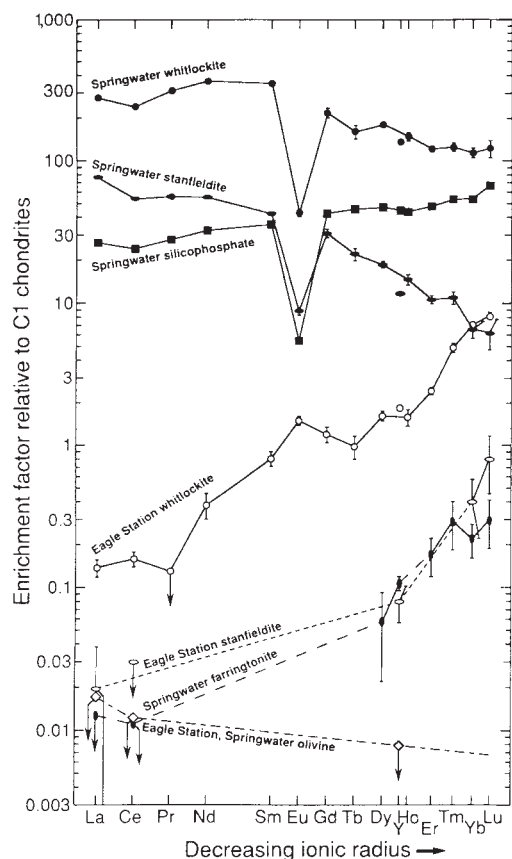


FIG. 1 Rare earth elements in olivine and phosphates in Eagle Station and Springwater were determined by ion microprobe. Error bars are $\pm 1\sigma$ and arrows indicate 2σ upper limits, based on counting statistics.

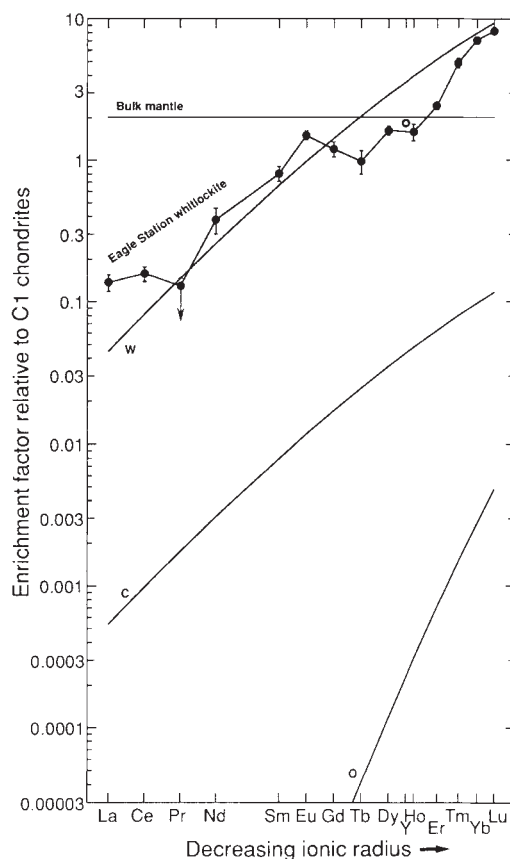


FIG. 2 A simple model was constructed to explain the REE pattern of Eagle Station whitlockite. A parent body of chondritic composition melted and a metallic core separated, leaving a mantle with a bulk REE enrichment factor of 2. Upon cooling, olivine crystallized and formed a cumulate, labelled 'C'. After metal and silicate mixed, they reacted to form phosphate-pyroxene rims on olivine. REE partitioned between 1.2% whitlockite, labelled 'W', and 98.8% olivine, labelled 'O'.

originally dissolved in molten metal¹³. Olivine is probably the source of calcium, as calcium solubility in olivine increases with temperature, and pallasitic olivine has calcium-depleted rims consistent with transport of calcium from olivine to calcium-rich phases¹⁴. Troilite and low-calcium pyroxene in the rims may form by reaction between dissolved sulphur in molten metal and the fayalite component of olivine. Troilite-enstatite rims on olivine clasts have been observed in a lunar breccia¹⁵.

If such reactions occurred, the phosphate minerals would concentrate the REE from olivine. Recent measurements of olivine/melt¹⁶ and whitlockite/melt¹⁷ partition coefficients for REE allow construction of a simple model to explain the REE patterns in Eagle Station. We assume that a chondritic parent body has totally melted, and a metallic core has separated. On cooling, olivine crystallized and sank to the bottom of the silicate mantle to form a cumulate. If the bulk mantle started out with twice the C1 chondritic concentration of REE, early-crystallizing olivine would have had the REE pattern labelled C in Fig. 2. Cumulate olivine then mixed with molten metal at the core-mantle boundary. On cooling, phosphate-pyroxene rims formed at olivine-metal interfaces by redox reactions, and the REE partitioned between whitlockite and olivine to give the REE patterns labelled W and O in Fig. 2. The ratio of olivine to whitlockite, 82, was chosen such that the whitlockite pattern matched the measured Eagle Station pattern as closely as possible. This ratio is lower than that of the bulk meteorite, but the amount of olivine that exchanged REE with whitlockite may have been limited by diffusion distance. Thus, the REE pattern of Eagle Station whitlockite preserves evidence of olivine-melt fractionation as the mantle of a parent body of chondritic composition solidified. Most other pallasites studied contain stanfieldite or whitlockite with fractionated, heavy REE-enriched patterns similar to those in Eagle Station, and can be modelled similarly.

Stanfieldite, whitlockite and silicophosphate in Springwater are much higher in REE and have very different REE patterns from those in Eagle Station. Farringtonite in Springwater is extremely low in REE; its crystal structure possibly prevents significant substitution of elements with large ionic radius. Stanfieldite, whitlockite and silicophosphate seem to have the patterns expected for liquidus phosphate, yet the Springwater parent-body mantle cannot have contained enough phosphorus to saturate the melt with phosphate when cumulate olivine was crystallizing. Whitlockite saturation experiments on lunar basalt composition¹⁸ show that several per cent P_2O_5 must be present for whitlockite to have cocrystallized with olivine from the low-silica, high-temperature melt. A second possibility is that the phosphate grew from a trapped liquid remaining from cumulate formation. As olivine is the first phase to crystallize from a chondritic mantle, and no other silicates are present in Springwater, a trapped liquid must have been relatively primitive in terms of its igneous fractionation history. One would expect the phosphates to have inherited the composition of the liquid and to have had a relatively flat REE pattern with no europium anomaly. Also, one would expect to find other components of the trapped liquid, for example silicate phases enriched in elements that are incompatible in olivine, such as calcium, aluminum and titanium. It is possible to combine the REE patterns of the three REE-rich phosphates to obtain a flat REE pattern (with the exception of europium), but it requires a large fraction to be silicophosphate, whereas the most abundant of the REE-rich phosphate phases is stanfieldite.

A third possibility is that late-stage crystallization on the Springwater parent body may have led to silicate liquid immiscibility and formation of a REE-rich, phosphorus-rich ferrobasaltic melt of high density and low viscosity. It has been suggested that this occurred on the Moon¹⁹. This liquid may have percolated downward, but it is difficult to see how it could have reached the core-mantle boundary. The magmatic-type REE patterns of phosphates in Springwater are difficult to reconcile

with origin at the core-mantle boundary and indicate that pallasites may have formed nearer the surface of the Springwater parent body. The Springwater parent body may have experienced incomplete core separation, or crustal and core-mantle boundary materials could have been brought together following collisional disruption. The negative europium anomaly in the phosphate REE patterns of Springwater suggests that the source from which the phosphate differentiated was depleted in europium, because there is no appropriate host for europium in Springwater. Santa Rosalia stanfieldite has a REE pattern similar to that of Springwater stanfieldite, indicating a similar origin. □

Received 26 April; accepted 16 September 1991.

1. Scott, E. R. D. *Geochim. cosmochim. Acta* **41**, 349–360 (1977).
2. Davis, A. M. thesis, Yale Univ. (1977).
3. Clayton, R. N. & Mayeda, T. K. *Earth planet. Sci. Lett.* **40**, 168–174 (1978).
4. Buseck, P. R. & Holdsworth, E. *Mineral. Mag.* **41**, 91–102 (1977).
5. Olsen, E. & Fredriksson, K. *Geochim. cosmochim. Acta* **30**, 459–470 (1966).
6. Buseck, P. R. *Geochim. cosmochim. Acta* **41**, 711–740 (1977).
7. Scott, E. R. D. *Geochim. cosmochim. Acta* **41**, 693–710 (1977).
8. Zinner, E. & Crozaz, G. *Int. J. Mass Spectrom. Ion Proc.* **69**, 17–38 (1986).
9. Fahey, A. J., Goswami, J. N., McKeegan, K. D. & Zinner, E. *Geochim. cosmochim. Acta* **51**, 329–350 (1987).
10. Crozaz, G. et al. *Earth planet. Sci. Lett.* **93**, 157–169 (1989).
11. Reed, S. J. B., Smith, D. G. W. & Long, J. V. P. *Nature* **306**, 172–173 (1983).
12. Crozaz, G. & McKay, G. *Earth planet. Sci. Lett.* **97**, 369–381 (1990).
13. Swisher, J. H. & Turkdogan, E. T. *Trans. met. Soc. AIME* **239**, 426–431 (1967).
14. Leitch, C. A., Steele, I. M., Hutcheon, I. D. & Smith, J. V. *Lunar planet. Sci. X*, 716–718 (1979).
15. Norman, M. D. *Proc. Lunar planet. Sci. Conf.* **12**, 235–251 (1982).
16. Colson, R. O., McKay, G. A. & Taylor, L. A. *Geochim. cosmochim. Acta* **52**, 539–553 (1988).
17. Dickinson, J. E. & Hess, P. C. *Lunar planet. Sci. XIV*, 158–159 (1983).
18. Dickinson, J. E. & Hess, P. C. *Lunar planet. Sci. XIII*, 172–173 (1982).
19. Neal, C. R. & Taylor, L. A. *Geochim. cosmochim. Acta* **53**, 529–541 (1989).

ACKNOWLEDGEMENTS. We thank R. N. Clayton, G. A. McKay and R. O. Sack for helpful discussions. This work was supported by NASA.

Incremental charging of a molecule at room temperature using the scanning tunnelling microscope

H. Nejoh

Aono Atomcraft Project, ERATO, JRDC, 5-9-9 Tohkohdai, Tsukuba, Ibaraki 300-26, Japan

ADDING electrons to, or extracting them from, a sample of microscopic dimensions is opposed by a charging energy that depends on the object's capacitance. The stepwise charging ('Coulomb staircase') that is expected as a result has been observed for metal films and droplets, using the scanning tunnelling microscope (STM) to inject the charge^{1–3}. But because of the very small charging energy, the effect was observable only at very low temperatures (~4 K). Here I present evidence for incremental charging of a single molecule (a liquid crystal) from STM studies conducted at room temperature. I observe quasiperiodic variations in the current-voltage curve which I attribute to discrete changes in the molecule's charge. The width of the steps and the asymmetry between positive and negative bias, however, remain to be explained.

Averin and Likharev¹ have treated the effect of single-electron tunnelling oscillations and the incremental charging effect or Coulomb staircase². The Coulomb staircase phenomenon has been observed experimentally^{3–5} and Tsukada *et al.*⁶ have analysed the effect using the semiclassical approach. Suppose a metal droplet is placed on a metal substrate, the surface of which is covered by a layer of metal oxide. When an STM tip is held above this metal droplet and the bias voltage is scanned without applying any feedback, a staircase-like current-voltage ($I-V$)

characteristic is obtained. This staircase is explained as follows: as the absolute value of the negative tip bias is increased, electrons cannot transfer from the tip to the metal droplet until the bias voltage is high enough to overcome the charging energy $e^2/2C$, where C is the capacitance of both tunnel junctions (assumed for simplicity to be the same). At bias voltages corresponding to energies between $e^2/2C$ and $3e^2/2C$, the charge on the droplet remains at -1 . In this bias range, the I - V curve has a plateau. When the bias increases beyond $3e^2/2C$, the current jumps steeply again (the incremental charging effect), and this corresponds to a mean droplet charge of -2 . In fact, because current flows, the charge on the droplet will fluctuate about a mean value, owing to successive tunnelling events¹⁰.

These staircases have been observed only at extremely low temperatures (~ 4 K), as the charging energy is so small, and the step-like increment must be free from the disturbance of thermal fluctuations. In contrast, when a molecule is used instead of a metal droplet, the size of the molecule is about one-tenth that of a metal droplet, so a larger charging energy can be expected. The incremental charging effect should then be observable even at room temperature.

Here I have used as a small particle the liquid crystal (LC) 4'-n-heptyl-4-cyanobiphenyl (7CB, where 7 is the number of carbons in the alkyl group). This compound shows a nematic LC phase near room temperature, and previous STM studies⁷ have shown that the molecules form ordered layers on graphite. STM images of the molecule can be observed for long periods, and this suggests that the molecules in the surface layer remain inert. The platinum (111) surface was polished to a roughness of $1\text{ }\mu\text{m}$, chemically etched in a solution of HNO_3 and HCl (ratio 1:3), rinsed in ultra-pure water and dried in air. A droplet of nematic 7CB was placed on the platinum surface immediately after drying. A commercial STM (Nanoscope II, Digital instruments) was used, with a tip of as-cut platinum-iridium wire. The I - V curves were taken at ambient pressure, and the temperature could be controlled within a certain range. The curves were obtained as follows: the bias voltage was scanned while the tip height was held constant. The acquisition cycle involved running the feedback for a period of time to establish the

set-point current, then turning off the feedback and ramping the bias voltage. At each of 2,000 points along the ramp, a bias modulation was added to the bias voltage and the current was measured. This process of modulating the bias voltage was repeated 16 times. The curve was obtained by repeating this cycle 40 times. An STM tip was positioned above a LC molecule on the platinum substrate. This system of STM tip, molecule and substrate behaves like a double tunnel junction.

Figure 1a shows the current and Fig. 1b the current derivative as a function of the bias voltage at a sample temperature of 21°C . The bias voltage is $-1,000$ mV and the setpoint current is 0.2 nA. Figure 1a clearly shows the staircase, and Fig. 1b shows the periodic peaks, which are separated by ~ 150 mV when the sample has a positive bias and ~ 200 mV when the sample has a negative bias. Results were obtained at a range of sampling periods from 200 to 300 ms, and in every case the peak separations were the same, indicating that the data do not contain artefacts due to sampling. The peak separations correspond to a thermal energy equivalent to $\sim 1,700^\circ\text{C}$ and are well above the thermal energy at room temperature.

To test the reproducibility of the result, I altered the bias voltage successively from -10 to -100 , $-1,000$ and $-5,000$ mV while keeping the set-point current constant at 0.2 nA. I also changed the set-point current from 0.2 to 20 nA while keeping the bias voltage constant at $-1,000$ mV. For all the bias and set-current conditions, periodic peaks were seen, and the peak separation was the same. Increasing the bias voltage and set-point current by two orders of magnitude corresponds to decreasing the tip-molecule distance by ~ 2 Å. Even when the tip separation is decreased by 2 Å, the tunnelling condition is still kept. For the incremental charging effect, the electron transition probability of the junction between tip and molecule has to be smaller than that of the junction between molecule and substrate³. As the gap between the tip and molecule is larger than that between molecule and substrate, the double junction consisting of tip, molecule and substrate satisfies this requirement³, although the exact tip-molecule distance is not known. This experiment shows that a gap variation of 2 Å does not affect the incremental charging effect condition.

The transition temperature of 7CB from the nematic to isotropic phase is 42.8°C . When the sample temperature was increased to 42.8°C , the staircase obtained (Fig. 2a) was less

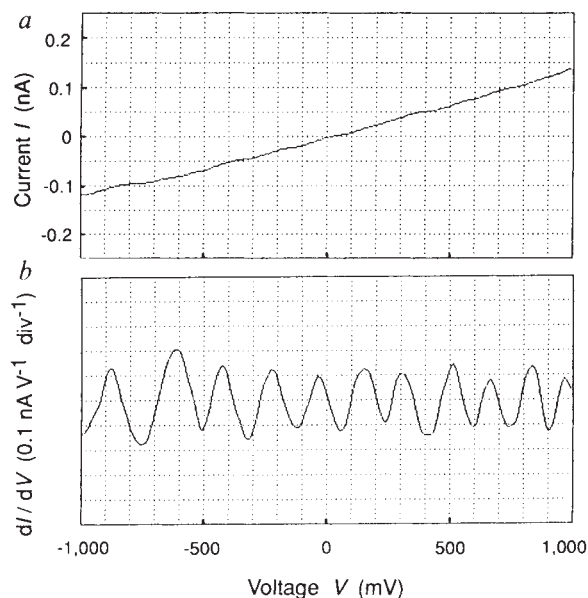


FIG. 1 a, Current as a function of bias voltage applied to a sample of 7CB on Pt(111), showing the 'staircase' effect. The sample temperature is 21°C . The bias voltage is $-1,000$ mV and the setpoint current is 0.2 nA. The bias modulation is 50 mV and the sample period is $250\text{ }\mu\text{s}$. b, Current derivative as a function of the bias voltage in a. The periodic peaks can be seen.

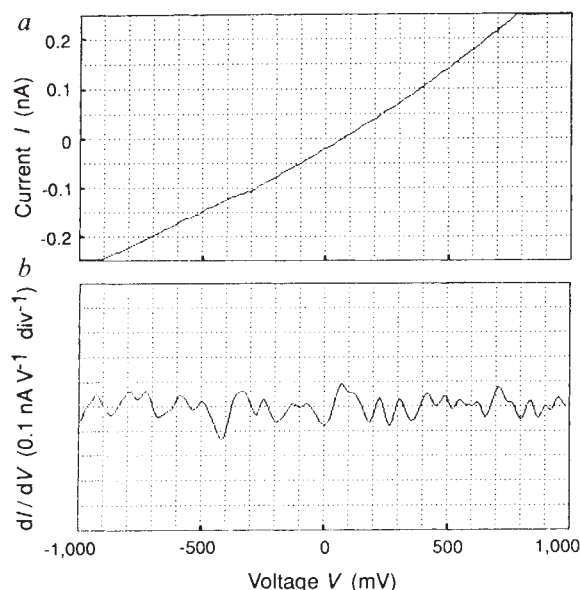


FIG. 2 a, The I - V curve at a sample temperature of 42.8°C . This is the transition temperature of 7CB from the nematic to the isotropic phase. All other conditions are as given in Fig. 1. b, Current derivative dI/dV against V for a.

distinct than that obtained at 21 °C. Figure 2b shows the periodic peaks, but the peak separation for positive sample bias changes from ~150 to ~70 mV and that for negative sample bias changes from ~200 to ~100 mV. When the sample temperature was increased to 100 °C, the current was more than that at 21 °C and 42.8 °C (Fig. 3a). This is probably due to the greater motion of LC molecules at higher temperature. As a result, the current through a molecule increases. At this temperature the staircase disappears (Fig. 3b). This suggests that the motion of the LC molecules in the isotropic phase hinders the staircase behaviour.

Peaks of the kind shown in Fig. 1 have been interpreted as evidence of the incremental charging effect by considering the capacitance of the system³⁻⁵. When a metal droplet of diameter 100 Å is used, the peaks are separated by ~10–50 mV, which corresponds to the charging energy. As the size of molecule used in my experiment is about one-tenth that of a conventional metal droplet, it seems that the peaks separated by ~100–200 mV in our experiment can also be interpreted on the basis of capacitance. But this is not the case, as I will discuss below.

First, I shall argue why the staircase cannot be attributed to the quantization of the energy levels of the LC molecule due to its finite size. A calculation of the energy levels for an isolated LC molecule⁸ (that is, where the interaction between molecule and substrate is ignored) shows that the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) are separated by ~10 eV. This indicates that the HOMO–LUMO gap is too large to produce the staircase observed here.

Next, I consider to what extent the I – V curves can be interpreted as proof of the incremental charging effect. The structure of the staircase expected theoretically would be perfectly periodic. The difference of the charging energy corresponding to the change from ne to $(n \pm 1)e$ is given by $\delta e^\pm = (n \pm 1)^2 e^2 / 2C - n^2 e^2 / 2C$ (refs 2, 6). The difference in the periodicity of the steps for positive and negative bias in my I – V curves shows that these results should be understood in terms of a new theory which is different from the capacitive charging energy treatment, the latter being applicable only to fairly large (100 Å) particles. The question is, what is the definition of the capacitance in the case of molecules? When the particle size is as little as 10 Å, the charging energy $n^2 e^2 / 2C$ could be replaced

by the total energy difference between the ionized state and the neutral state. The terrace widths of the staircase would then be different for ionized states, as the energies of the different ionized states will not in general be the same. According to this interpretation, one can gain insight into the different ionized states from the staircase behaviour.

A self-consistent-field calculation for an isolated cyanobenzene molecule shows that the total energy difference between the neutral state and the -1 ionized state is 1.96 eV, and that between the $+1$ and neutral state is 8.96 eV. Clearly these total energy differences are too large to explain the observed peak separation, although qualitatively the difference between negative and positive ionized states is in the same direction as that observed. But again, this calculation ignores electron correlations, whereas the peaks observed show that the total energy of a molecule is changed considerably when the molecule interacts with the substrate surface. This suggests that there might be strong electron correlation, which might therefore have to be included for a complete understanding of the experimental results (including the changes in peak separation with temperature).

These observations of incremental charging suggest the possibility of molecular devices based on single-electron tunnelling⁹ at room temperature. □

Received 7 May; accepted 13 August 1991.

1. Averin, D. V. & Likharev, K. K. *J. Low Temp. Phys.* **62**, 345–373 (1986).
2. Likharev, K. K. *IBM J. Res. Develop.* **32**, 144–158 (1988).
3. Mullen, K. & Ben-Jacob, E. *Phys. Rev. B* **37**, 98–105 (1988).
4. van Bentum, P. J. M., Smokers, R. T. M. & van Kempen, H. *Phys. Rev. Lett.* **60**, 2543–2546 (1988).
5. Wilkins, R., Ben-Jacob, E. & Jaklevic, R. C. *Phys. Rev. Lett.* **63**, 801–804 (1989).
6. Tsukada, M., Shima, N., Kobayashi, K., Inada, K. & Mizokami, T. *Prog. theor. Phys. Suppl.* **101**, 221–230 (1990).
7. Smith, D. P. E., Horber, J. K. H., Binnig, G. & Nejjoh, H. *Nature* **344**, 641–644 (1990).
8. Nejjoh, H. *Appl. Phys. Lett.* **57**, 2907–2909 (1990).
9. Likharev, K. K. *IEEE Trans. Mag.* **23**, 2, 1142–1145 (1987).
10. Veda, M. *Phys. Rev. B* **42**, 3087–3099 (1990).

ACKNOWLEDGEMENTS. I thank M. Tsukada, S. Watanabe and M. Aono for useful discussion.

Synthesis of pentacoordinate silicon complexes from SiO₂

Richard M. Laine^{*†}, Kay Youngdahl Blohowiak^{*}, Timothy R. Robinson^{*}, Martin L. Hoppe^{*}, Paola Nardi^{*}, Jeffrey Kampf[†] & Jackie Uhm^{*}

^{*} Department of Materials Science and Engineering, University of Washington, Seattle, Washington, USA

[†] Departments of Materials Science and Engineering, and Chemistry, University of Michigan, Ann Arbor, Michigan 48109-2136, USA

THE potential role of inorganic and organometallic silicon compounds in the development of new chemical reagents, polymers, glasses and ceramics¹ is limited at present by the paucity of simple silicon-containing starting materials. Whereas industrial carbon-based chemistry can draw on the diversity of compounds produced from crude oil, coal or other natural sources, silicon chemistry² relies almost exclusively on the carbothermal reduction of SiO₂ to silicon. This is then transformed into feedstock chemicals by reaction with HCl, or by routes such as the 'direct process' for making methylchlorosilanes², in which silicon is reacted with methyl chloride at 200–350 °C over a copper/tin catalyst. Organosilicon compounds are in demand in fields ranging from organic synthesis to ceramics to the electronics industry. New synthetic routes to these materials are therefore highly desirable, especially if they rely on low-cost SiO₂ and on processing methods that avoid the energy-intensive and equipment-intensive carbothermal reduction step which currently precedes almost all silicon

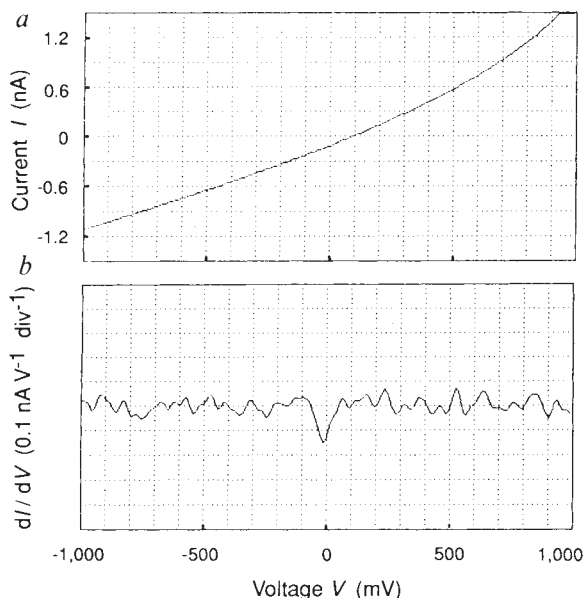
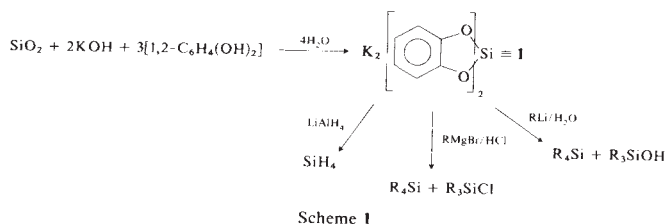


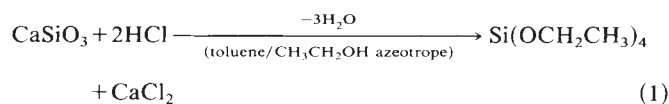
FIG. 3 *a*, The I – V curve at a sample temperature of 100 °C. The staircase cannot be seen, and the current is higher than for results obtained at lower temperatures. *b*, Current derivative dI/dV against V for *a*.

chemistry. Here we describe a direct process in which SiO_2 is reacted with ethylene glycol and an alkali base to produce highly reactive, pentacoordinate silicates which provide access to a wide variety of new silicon compounds.

Earlier investigations³⁻⁵ have explored the chemical reactivity of the hexacoordinate silicon compound tris(catecholato) silicate, **1**, readily made by reaction of silica, sand or even quartz with catechol (1,2-dihydroxybenzene) in basic media (see scheme 1)⁶⁻⁸. Unfortunately, **1** is quite stable and can only be modified usefully by reaction with strong nucleophiles:



The work of Kenney and Goodwin⁹, which is complementary to the work in refs 3-5, demonstrates that protonation of mineral silicates followed by careful azeotropic removal of water provides up to 70% yields of the tetracoordinate silicon compound, $\text{Si}(\text{OCH}_2\text{CH}_3)_4$:



We describe here a general method of synthesizing pentacoordinate, rather than hexa- or tetra-coordinate silicates, directly from SiO_2 , ethylene glycol and base. The resulting glycolato silicates are very reactive, inexpensive and offer unique opportunities for the synthesis of a wide variety of silicon containing chemicals and polymers. Furthermore, they offer the opportunity to develop new routes to silicon-containing glasses, ceramics and zeolites.

A mixture of 60 g silica gel ($>600 \text{ m}^2 \text{ g}^{-1}$), fused silica (325 mesh , $0.5\text{--}0.8 \text{ m}^2 \text{ g}^{-1}$) or sand ($<0.2 \text{ m}^2 \text{ g}^{-1}$), 1.1 equivalents of a group 1 metal hydroxide (for example 44 g KOH) and an excess (2 l) of ethylene glycol (or other 1,2-diol, such as propane-1,2-diol) was heated under N_2 in a magnetically stirred, standard pyrex distillation setup so that the ethylene glycol was distilled off slowly (with removal of water), to result in the dissolution of the silica gel (1-2 h), fused silica (3-6 h) or sand ($>200 \text{ h}$).

The products were isolated by cooling the much-reduced volume of the reaction solution to effect crystallization, or by precipitation with CH_3CN . Further washing with CH_3CN , followed by drying under vacuum (130°C), provided white powders in $>80\%$ isolated yields (350 g where KOH was used). The reaction can be run using Li, Na, K or Cs hydroxide.

The products are essentially insoluble in all polar solvents except CH_3OH (but not in $\text{CH}_3\text{CH}_2\text{OH}$). Recrystallization is readily effected using $\text{CH}_3\text{OH}/\text{CH}_3\text{CN}$. The following results are from a typical elemental analysis for $\text{M} = \text{K}$ (analysis carried out by Galbraith Laboratories, Inc., Knoxville) calc. (found) 27.53 (27.63)% C; 4.98 (4.64)% H; 13.60 (12.92)% Si; 17.84 (17.99)% K; 37.01 (36.81)% O (by difference). These results indicate a dimeric pentacoordinate species, labelled **2** in reaction (2) (refs 10, 11). In this scheme M is Li, Na or K. The Cs salt can be isolated as a monomer. Analysis gave typical values calculated (found) as 20.72 (21.06)% C; 3.63 (3.83)% H; 8.58 (8.21)% Si; 39.38 (38.84)% Cs; 27.32 (27.06)% O by difference. This indicates that a monomeric species, such as **3** in scheme 2 (below) was formed as

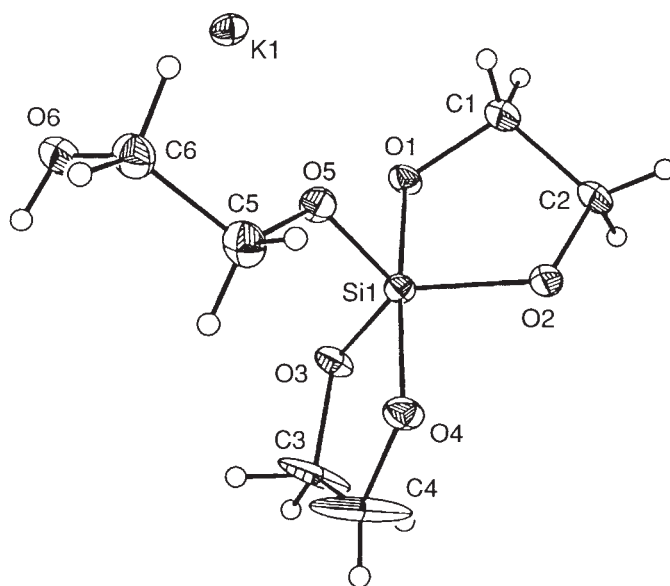


FIG. 1 X-ray single crystal structure of $\text{KSi}(\text{OCH}_2\text{CH}_2\text{O})_2\text{OCH}_2\text{CH}_2\text{OH}$. Crystal grown from ethylene glycol/acetonitrile.

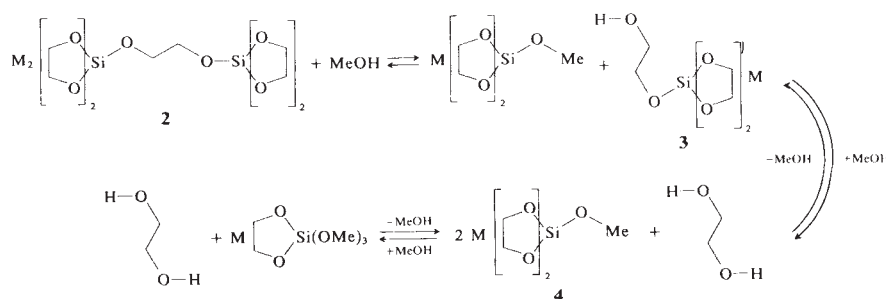
supported by the infrared (nujol) peak found for O-H at $3,300 \text{ cm}^{-1}$ and the solubility in dimethylsulphoxide (DMSO).

NMR studies suggest that the products dissolve in CD_3OD through an exchange reaction. The ^{13}C spectra for all of the salts (Li, Na, K and Cs) show two peaks at 61.3 ± 0.3 and $64.3 \pm 0.3 \text{ p.p.m.}$ Similarly, a single peak appears in the ^{29}Si NMR at $\sim -103.1 \pm 0.3 \text{ p.p.m.}$ The ^1H spectra all contain one broad singlet at $\sim 3.4 \text{ p.p.m.}$ The ^{13}C peak at 64.3 p.p.m. and the ^1H peak at 3.4 p.p.m. are consistent with free ethylene glycol, despite analytical results that indicate the materials are pure. The ^{29}Si NMR peaks are consistent with those previously reported for aliphatic, pentacoordinate silicates¹¹⁻¹⁴.

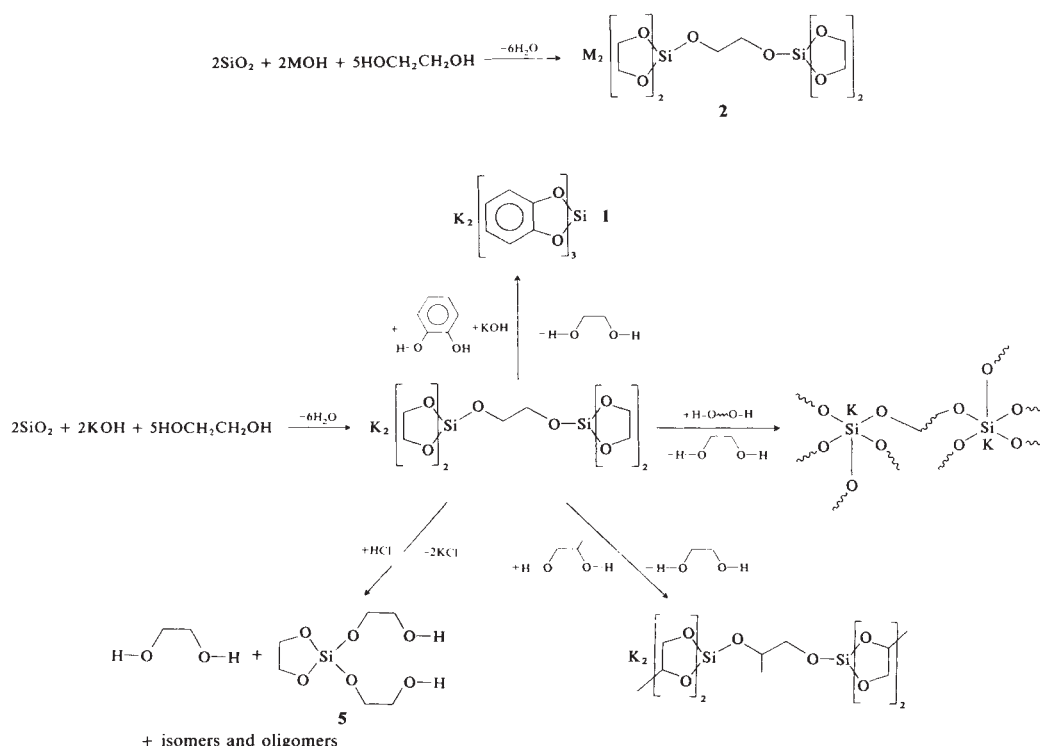
A plot of the ratio of the ^{13}C peak heights for free ethylene glycol to those for bound glycol, as a function of temperature for the dry K salt dissolved CD_3OD , shows a first-order increase in free ethylene glycol as the temperature is lowered to -50°C . At -50°C , the ratio of free to bound glycol is $0.85 (\pm 0.05)$. Therefore, it appears that exchange proceeds beyond replacement of a bridging or monodentate glycol (in the Cs salt) with more than one CD_3O^- group. The likely exchanges are shown in scheme 2.

The extent of the exchanges and the exact mechanism(s) whereby they occur must await more detailed kinetic studies.

Further support for the alcohol exchange process and for an intermediate monomeric species comes from the recrystallization of the dimer from ethylene glycol/acetonitrile which provides X-ray-quality crystals. The structure shown in Fig. 1 was refined to $R_w = 0.0492$. The structure is typical of an ionic solid in that the potassium contact distances between nearest neighbours are essentially identical. As is common for pentacoordinate, aliphatic silicate complexes, the ligand oxygens



Scheme 2



occupy a trigonal bipyramidal geometry about the silicon¹²⁻¹⁴.

The ready exchange of alkoxy ligands permits the synthesis of a wide variety of derivatives. Refluxing $M_2[Si_2(OCH_2CH_2O)_5]$ in an excess of a different 1,2-diol (for example, 1,2-propanediol or pinacol (2,3-dimethyl-2,3-butanediol)) followed by a removal by distillation of excess diol and free ethylene glycol (0.5–2 h at the reaction temperature) results in complete ligand exchange and the formation of new dimeric materials (following drying under vacuum at 130 °C), as illustrated in scheme 3. Pinacolate derivatives have been prepared previously¹²⁻¹⁴. Reaction of **2** with 6 equivalents of catechol and an additional equivalent of base provides the tris(catecholato)Si²⁻ species of scheme 1, in quantitative yield. With 4–5 equivalents and no additional base, the tris species is also isolated, but novel polymers are produced simultaneously (ref. 15, and R.M.L. and N. Budrys, unpublished results).

Exchange with other diols including 1,3-propanediol, $H(OCH_2CH_2)_4OH$ and hydroquinone produces polymer products with anionic silicon moieties in the polymer backbone (scheme 3). Detailed characterization of the polymers will be reported elsewhere, but the following example illustrates the simplicity of the process.

Heating a mixture of 9.9 g (24 mmol) of $Li_2[Si_2(OCH_2CH_2O)_5]$ and 30 g (124 mmol) of $H(OCH_2CH_2)_4OH$ under vacuum (~0.1 mm Hg) for ~3 h at 130 °C results in dissolution

of the lithium salt and recovery of 8–12 ml of a clear liquid distillate (depending on total heating time) that consists of roughly equal molar amounts of $H(OCH_2CH_2)_4OH$ and $HOCH_2CH_2OH$ (determined by gas chromatography). The material remaining in the flask becomes increasingly viscous with time as bidentate glycol groups are replaced by monodentate, crosslinking $H(OCH_2CH_2)_4OH$. The new polymer is soluble in CH_3OH , CH_3CH_2OH and 1:1 $CH_3CH_2OH:CH_3CN$ and may be useful as an ion conductor for battery applications^{16,17}.

Finally, pyrolysis of the glycolato silicates leads directly to silicate glasses or glass ceramics¹⁹. The alkali complexes, when pyrolysed in air at 10 °C min⁻¹ to temperatures above 300 °C for at least 2 h, provide the line compounds $M_2Si_2O_5$ (ref. 18). If excess base is used, then partially crystallized alkali-rich silicate glasses are formed preferentially.

The four-coordinate alkoxy silane complex, **5**, obtained by neutralization of **2** with HCl gas, can be used as a substitute for $Si(OCH_2CH_3)_4$ in sol-gel reactions. Unlike $Si(OCH_2CH_3)_4$, **5** is miscible with H_2O , offers rheologically useful properties for coating and fibre applications and can be used directly to form glass-polymer composites, 'ceramers'. It can also be mixed with the various alkali derivatives of **2** and heated to form alkali-poor silicate glasses (unpublished results, R.M.L. *et al.*). □

Received 19 March; accepted 27 August 1991.

1. *Inorganic and Organometallic Polymers with Special Properties* (ed. Laine, R. M.) NATO ASI Series E: Appl. Sci., No. 30 (Kluwer, Dordrecht, 1991).
2. *Kirk-Othmer Encyclopedia of Chemical Technology*, 3rd Ed., Vol. 20, 750–880 (Wiley-Interscience, New York, 1979).
3. Boudin, A., Cerveau, G., Chuit, C., Corriu, R. J. P. & Reye, C. *Angew. Chem. Int. Ed.* **25**, 473–474 (1986).
4. Corriu, R. J. P., Perz, R. & Reye, C. *Tetrahedron* **39**, 999 (1983).
5. Boudin, A., Cerveau, G., Chuit, C., Corriu, R. J. P. & Reye, C. *Organomet.* **7**, 1165–1171 (1988).
6. Rosenheim, A., Raibmann, B. & Schendel, G. Z. *anorg. Chem.* **196**, 160 (1931).
7. Barnum, D. W. *Inorg. Chem.* **9**, 1942 (1970).
8. Barnum, D. W. *Inorg. Chem.* **11**, 1424 (1970).
9. Kenney, M. E. & Goodwin, G. B. *US Patent No.* 4,717,773 (January 1988).
10. Frye, C. L. *J. Am. chem. Soc.* **92**, 1205–1210 (1970).
11. Tandura, S. N., Voronkov, M. G. & Alekseev, N. V. *Topics Curr. Chem.* **131**, 99–186 (1986).
12. Swamy, K. C. K. *et al. J. Am. chem. Soc.* **112**, 2341–2348 (1990).

13. Holmes, R. R. *Chem. Rev.* **90**, 17 (1990).
14. Holmes, R. R., Day, R. O. & Payne, J. S. *Phos. Sulf. Silicon* **42**, 1–13 (1989).
15. Laine, R. M., Ray, D. J., Viney, C. & Robinson, T. R. *Am. chem. Soc. Polym. Preprints* **32**, No. 3, 550 (1991).
16. Fujita, M. & Honda, K. *Polymer Commun.* **30**, 200 (1989).
17. Spindler, R. & Shriver, D. F. *J. Am. chem. Soc.* **110**, 3036 (1988).
18. Levin, E. M., Robbins, C. R. & McMurdie, H. F. *Phase Diagrams for Ceramists*, 5th Printing (eds. Reser, M. K.) 87–92 (American Ceramic Society, 1985).
19. Bickmore, C. R. *et al. 5th Int. Conf. Ultrastructure Processing* (eds. Hench, L. L., West, J. K. & Ulrich, D. R.) (Wiley, New York, in the press).

ACKNOWLEDGEMENTS. We thank J. Negrych and Minco, Inc. for discussions and gifts of fused silica and virgin sand. We also thank R. J. P. Corriu and P. Knochel for discussions and suggestions. We thank B. Posin for early contributions. We thank the Air Force Office of Scientific Research and the Office of Naval Research for support.

Generation of Maxwell displacement current across an azobenzene monolayer by photoisomerization

Mitsumasa Iwamoto*, Yutaka Majima,
Haruhiko Naruse, Tetsuya Noguchi & Hiromasa Fuwa

Department of Physical Electronics, Tokyo Institute of Technology, 2-12-1,
O-okayama, Meguro-ku, Tokyo 152, Japan

* To whom correspondence should be addressed

INTEREST in the behaviour of Langmuir–Blodgett (LB) films¹ is motivated by possible applications that include molecular electronics, biosensors and modelling of biological membranes. We recently^{2–5} reported an electrical technique that allows molecular motions in LB films to be probed by measuring the Maxwell displacement current⁶ generated across electrodes above and below the film: this current results from changes in the orientation of the molecular dipoles. Here we use the technique to study the effect of photoisomerization in an LB film consisting of an azobenzene derivative, which can be switched photochemically between *cis* and *trans* isomers. Films of this sort are of particular interest because of their possible use in information-storage devices^{7–11}. Successive isomerizations induced by irradiation with ultraviolet and visible light produced transient displacement-current pulses between the electrodes. We anticipate that this technique will become widely used for dynamical studies of LB films.

According to Maxwell's electromagnetic field theory⁶, the total current flowing across a monolayer is the sum of the conduction current and displacement current. The conduction current is the steady-state current, and it is generated by applying a potential difference between two electrodes which are separated by the monolayer; electrons injected from one electrode into the monolayer are conveyed through it to the counter electrode under the external electric field thus produced. Both top and base electrodes must be brought into contact with a single monolayer without destroying it to detect the conduction current. This is, unfortunately, very difficult, possibly because of unexpected defects originating in a single monolayer deposited on the base electrode¹² and destruction of the monolayer by the application of the top electrode¹³. In contrast, the displacement current is the transient current, and it is generated only when the electric flux formed between two electrodes changes with time because an external stimulus (for example

light, pressure or heat) is applied to the monolayer². As the electric flux diverging from charged particles in a single monolayer falls on electrodes even when they are suspended in air, the electrodes need not be in contact with the monolayer. The displacement current measurement can therefore detect the motion of charged particles, such as electron transfer or orientational change in polar molecules, in the planned artificial monolayers without destroying the monolayers.

Techniques used to detect changes in the physicochemical properties of artificial mono- and multilayer systems have recently received much attention. For example, the photochromic properties of azobenzene derivatives have been established for the development of optical memories^{7–11}: *trans*-to-*cis* isomerization occurs when they are irradiated with ultraviolet light, and the reverse occurs at visible wavelengths. Unfortunately, to our knowledge, no electrical technique has been established that can detect the motion of charged particles induced by photo-isomerization in a single monolayer, although the photoelectric response of a bilayer of a lipid membrane containing an azobenzene derivative has been reported¹⁴. Detection of the photoelectric response of single monolayers might be essential for a better understanding of the relationship between the structure and function of monolayers. For the past few years, we have been developing a displacement current technique used to analyse the dynamic motion of monolayers on a water surface, where the motion was generated by applying surface pressures^{2–5}. Here we use the technique to detect the displacement current generated by irradiation of an azobenzene monolayer both on water and on a solid substrate.

Figure 1 shows our experimental arrangement. Electrode 1, composed of a transparent SnO₂-coated glass slide, was suspended in air, parallel to the water surface, whereas electrode 2 was immersed in the water subphase. The two electrodes were connected by a gold wire through a sensitive electrometer, whose internal electrical resistance was found to be zero. A 500-W xenon lamp was used to induce photoisomerization in a single monolayer. The ultraviolet and visible light outputs of the lamp were separated using a glass filter. The azobenzene used here was AZBPAA, which was delivered onto the surface of deionized water from a benzene solution to form a single monolayer:

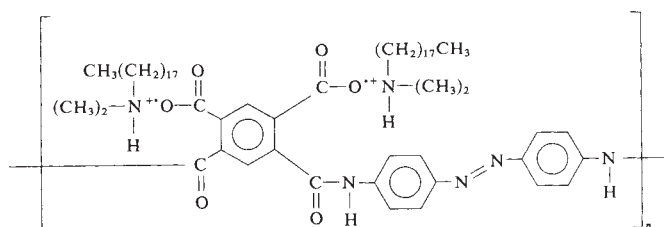
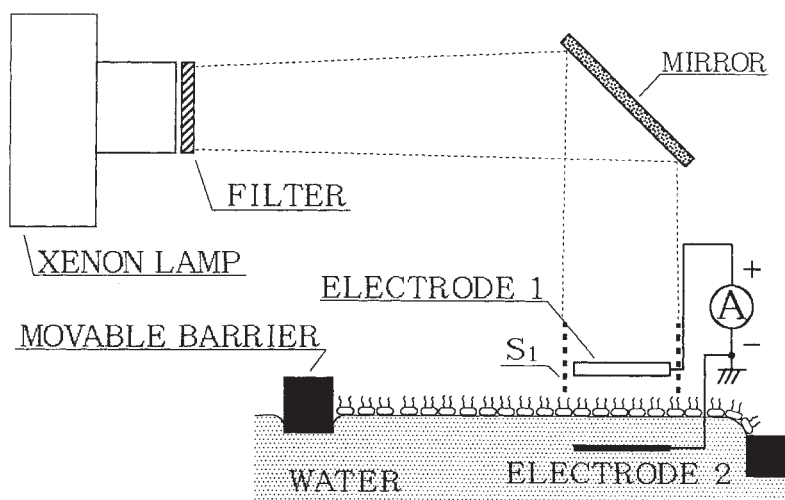


FIG. 1 The experimental arrangement. The working area of electrode 1 was 80 cm². The spacing between electrode 1 and the water surface was adjusted to be 1.0 mm with the help of a micrometer, as described in ref. 5. The LB trough was filled with deionized water (pH 5.8) at 20.0 °C. S₁ represents the copper plate used to shield electrode 1.



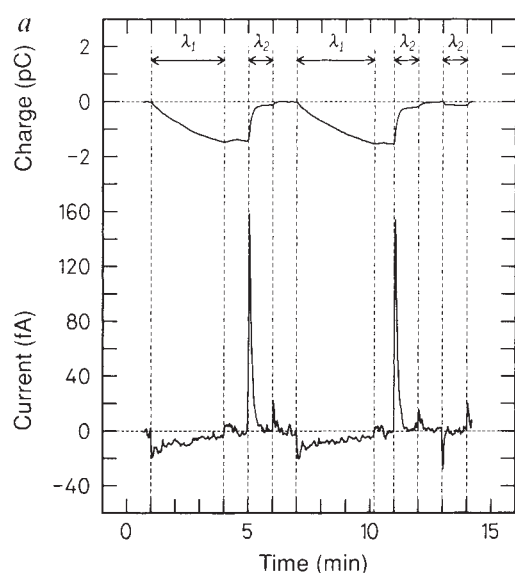
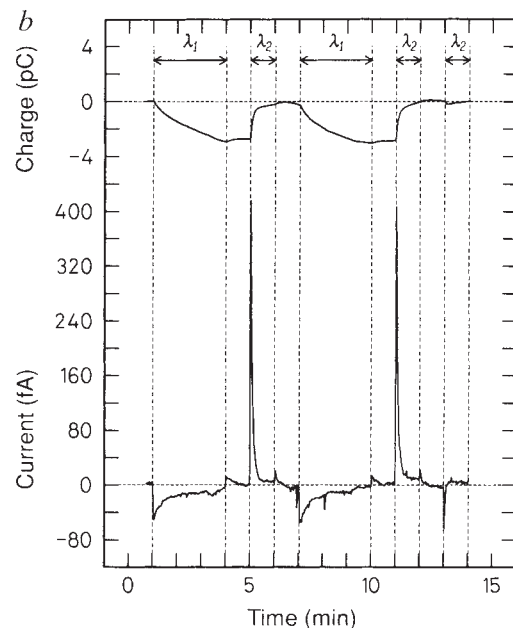


FIG. 2 *a*, Generation of displacement current from a single monolayer formed on the surface of water. The current was generated by alternate irradiation with ultraviolet (λ_1) and visible (λ_2) light and subsequently by two successive applications of visible (λ_2) light. The charge flowing through the circuit during the photoirradiation is also shown. The experiment was performed at a surface pressure of 45 mN m^{-1} ; that is, the area per molecule was $60 \text{ \AA}^2$. *b*, Generation of displacement current from a single monolayer deposited on a glass slide coated in SnO_2 . The current was generated by irradiation



with ultraviolet light and flows from the base SnO_2 electrode to the top SnO_2 electrode through an electrometer. The charge flowing through the circuit during the photoirradiation is also shown. Displacement current was never generated from a device with a $\text{SnO}_2/(\text{air-gap})/\text{SnO}_2$ structure.

Most of the AZBPAA molecules were found to be in the *trans* form when they were spread on the water surface. The area of the LB trough to be covered by the monolayer was controlled with the help of the movable barrier. A Wilhelmy balance consisting of a filter paper was used to measure surface pressure. As described previously⁵, displacement current, which is proportional to the capacitance C_1 formed between electrode 1 and the water surface, flows only because of changes in the orientation of polar molecules on the water surface, changes in the density of molecules distributed on the water surface and changes in the surface potential of the water. One or more of these changes might generate displacement current when a single monolayer on the water surface is irradiated through the transparent electrode 1. If electrode 1 is electrically shielded by a copper plate (Fig. 1), we can measure the displacement current flowing through the circuit with high sensitivity and high accuracy¹⁵. The photochromic properties of AZBPAA monolayers formed on the water surface have recently been established; *trans*-to-*cis* and *cis*-to-*trans* photo-induced isomerization in AZBPAA monolayers have been found to occur on irradiation with ultraviolet light centred at 380 nm (λ_1) and with visible light centred at 450 nm (λ_2), respectively. For our measurements of displacement current, therefore, we applied ultraviolet light at wavelength λ_1 and visible light of wavelength λ_2 .

Figure 2*a* shows a typical example of the displacement current obtained by irradiation alternately by light at λ_1 and λ_2 (repeated twice), followed by radiation with visible light λ_2 . The alternate irradiations generated a displacement current flowing from electrode 2 to electrode 1 when ultraviolet irradiation (λ_1) was applied and in a current in the opposite direction when visible light (λ_2) was applied. The total charge flowing through the circuit during irradiation with the ultraviolet light was almost the same as that flowing during irradiation with the visible light. These results indicate that the dynamic motion of molecules induced by the *cis*-*trans* photoisomerization is reversible. On the other hand, irradiation twice in succession with visible or ultraviolet light (not shown) never generated the displacement

current. These findings suggest that our technique satisfies the requirement for the construction of monolayer memory devices. Interconversion between two states the monolayers can easily be detected from the direction of the current in Fig. 2*a*, although the conversion between the two states was estimated to be $\sim 10\%$ from calculations based on difference spectra¹⁶.

To clarify the physical mechanisms relating to the displacement current generation, we deposited a monolayer of *trans*-AZBPAA on a flat, transparent SnO_2 -coated glass slide by the conventional LB technique, where the surface pressure and subphase temperature during film fabrication were controlled at an appropriate pressure and 20°C , respectively. We then fabricated devices incorporating an air gap, with a working area of 8.75 cm^2 ($2.5 \times 3.5 \text{ cm}$) and a spacing of $39.4 \text{ }\mu\text{m}$, to form a

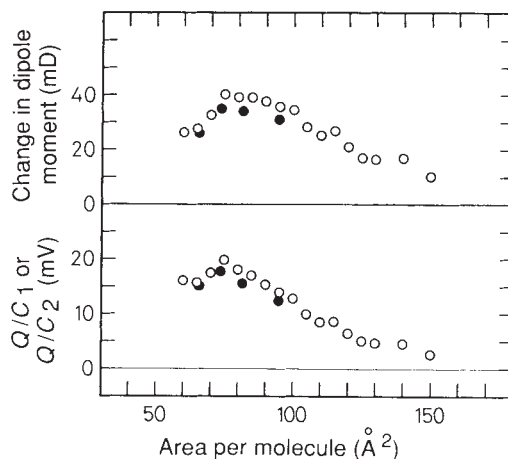


FIG. 3 Plots of the charge Q/C_1 (○) generated from a single monolayer formed on the water surface and of the charge Q/C_2 (●) generated from a single monolayer deposited on solid substrates during irradiation with visible light (λ_2). The change in the vertical component of dipole moment is also plotted.

structure consisting of base electrode (coated with SnO_2), monolayer, air-gap and top SnO_2 electrode as in Fig. 1. We then measured the displacement current produced by alternative applications of ultraviolet light λ_1 and visible light λ_2 . We also measured the current on successive irradiation with visible or ultraviolet light. Displacement current was generated (Fig. 2b) in a similar manner to that generated on the water surface, whereas it was never initiated for devices without a monolayer. As the molecules deposited on solid SnO_2 slides cannot move, there is no possibility that the current generated is due to a change in the density of molecules distributed on the slides. Therefore, we conclude that orientational change caused by the photoinduced *cis-trans* isomerization of molecules is responsible for the current. It is interesting to note that if the charge, Q , flowing through the circuit during irradiation with visible light is divided by the capacitance, C , of the device used, Q/C_2 for monolayers deposited on solid substrates is almost the same as Q/C_1 for monolayers on the surface of water, as plotted in Fig. 3. On water and on solid substrates, irradiation-induced changes in the vertical component of average dipole moment of the constituent monolayer molecules therefore seem to be almost the same (Fig. 3). Thus, we conclude that neither changes in the surface potential nor changes in the density of molecules distributed on the water surface make a significant contribution to the displacement current generated by the photoisomerization of AZBPAA monolayers on water; that is, orientational change in the molecules is the main contribution on water. $\square$

Received 28 January; accepted 14 August 1991.

- Langmuir, I. *J. Am. chem. Soc.* **39**, 1848 (1917).
- Iwamoto, M. & Majima, Y. *Japan J. appl. Phys.* **27**, 721–725 (1988).
- Iwamoto, M. & Majima, Y. *Thin Solid Films* **178**, 67–72 (1989).
- Majima, Y. & Iwamoto, M. *Jap. J. appl. Phys.* **29**, 564–568 (1990).
- Iwamoto, M. & Majima, Y. *J. chem. Phys.* **94**, 5135–5142 (1991).
- Collin, R. E. in *Field Theory of Guided Waves*, 1–5 (McGraw-Hill, New York, 1960).
- Yabe, A. et al. *Thin Solid Films* **160**, 33–41 (1988).
- Takahiro, S. & Ichimura, K. *Thin Solid Films* **179**, 77–83 (1989).
- Liu, Z. F., Loo, B. H., Baba, R. & Fujishima, A. *Chem. Lett.* 1023–1026 (1990).
- Kawai, T., Uemura, J. & Takenaka, T. *Langmuir* **5**, 1378–1383 (1989).
- Liu, Z. F., Hashimoto, K. & Fujishima, K. *Nature* **347**, 658–660 (1990).
- Peterson, I. R. *J. molec. Electron.* **2**, 95–99 (1986).
- Couch, N. R., Montgomery, C. N. & Jones, R. *Thin Solid Films* **135**, 173–182 (1986).
- Fujiwara, H. & Yonezawa, Y. *Nature* **351**, 724–726 (1991).
- Iwamoto, M. & Majima, Y. *Jpn. J. appl. Phys.* **28**, 1922–1925 (1989).
- Fischer, E. *J. phys. Chem.* **71**, 3704–3706 (1967).

ACKNOWLEDGMENTS. We thank M. Kakimoto and Y. Imai of Tokyo Institute of Technology, and K. Iriyama of the Jikei University School of Medicine.

A stable-isotope tree-ring timescale of the Late Glacial/Holocene boundary

Bernd Becker*, Bernd Kromer† & Peter Trimborn‡

* Institut für Botanik, Universität Stuttgart-Hohenheim 210, 7000 Stuttgart, Germany

† Institut für Umweltphysik, Universität Heidelberg, INF 366, 6900 Heidelberg, Germany

‡ GSF Institut für Hydrologie, 8042 Neuherberg, Germany

LATE Glacial and Holocene tree-ring chronologies, like deep-sea sediments or polar ice cores, contain information about past environments. Changes in tree-ring growth rates can be related to past climate anomalies and changes in the isotope composition of tree-ring cellulose reflect changes in the composition of the atmosphere and the hydrosphere. We have established a 9,928-year absolutely dated dendrochronological record of Holocene oak (*Quercus robur*, *Quercus petraea*)—and a 1,604-year floating Late Glacial and Early Holocene chronology of pine (*Pinus sylvestris*) from subfossil tree remnants deposited in alluvial terraces of south

central European rivers. The pine sequence provides records of dendro-dated ^{14}C , ^{13}C and ^2H patterns for the late Younger Dryas and the entire Preboreal (10,100–9,000 yr BP). Through the use of dendrochronology, radiocarbon age calibration and stable isotope analysis, we suggest that the Late Glacial/Holocene transition may be identified and dated by ^{13}C and ^2H tree-ring chronologies.

Through the stepwise cross-dating of more than 6,000 tree-ring curves, we established the absolute German oak dendrochronology, continuous to the year 7938 BC. It consists of subfossil river oaks linked with prehistoric, historic and modern oak samples. The statistical significance of the replication of the German oak chronology from the present to 4000 BC is already described in detail^{1–3}. The master chronology from before 4000 BC consists of an unbroken record of subfossil trees from the Main river. It is confirmed by chronologies of the Danube and Rhine rivers⁴.

This unique tree-ring calendar provides precise dendrochronological dating of prehistoric oak wood constructions from central Europe¹. It also allows for calibration of radiocarbon ages beyond the hitherto longest Holocene tree-ring records of US Bristlecone pine and Irish oak^{4,5}. The oldest Holocene oak trunks, not yet part of the absolute chronology, date from between 9,300 to 9,000 ^{14}C yr BP. Inclusion of this material through additional cross-matching has the potential of adding another 300 dendro-years to the absolute oak chronology.

A further extension of the European tree-ring calendar can be achieved by using remnants of pines from Early Holocene and Late Glacial valley deposits. We have cross-matched 183 pines to a 1,604-year-long floating pine masterchronology. As will be shown below, this chronology begins at the late Younger Dryas and ends with the beginning of the absolute oak sequence at the early Boreal. It must therefore cover the Pleistocene/Holocene boundary. The oldest radiocarbon-dated section of the pine chronology has a ^{14}C age of $10,030 \pm 30$ ^{14}C yr BP and the youngest one (ring number 1,510) is $8,930 \pm 30$ ^{14}C yr BP. (For absolute ages, the term 'dendro-years BP' refers to tree-ring ages before 1950 AD; for conventional radiocarbon ages, the term ^{14}C BP is used.) For its absolute dating, the position with respect to the oak chronology must be known. The absolute placement of the floating pine series is derived by the following procedure: from the patterns of ^{14}C ages versus dendro-years at the end of the Preboreal pine and at the beginning of the oak chronology, a characteristic (century scale⁶) variation of the ^{14}C level can be deduced. This observation is based on both the rapid decrease of ^{14}C ages between 9,100 and 8,930 ^{14}C yr BP at the end of the pine series (caused by increasing ^{14}C levels) and a 350 dendro-year long plateau of roughly constant ^{14}C ages at the beginning of the oak chronology (Fig. 1b).

Similar ^{14}C oscillations have been found over most of the Holocene calibration curve⁶. They are probably caused by changes in solar activity and correspond to a 2–4% increase in atmospheric ^{14}C superimposed on the Holocene long-term trend. The floating pine series is adjusted to the absolute oak series to create a ^{14}C oscillation of this type (Fig. 1a; oscillation C1/C2).

The characteristic shape of the ^{14}C oscillation at the link of both sets narrows constraints to the absolute placement of the pine series: a younger absolute age would lead to an overlap of the measured radiocarbon sections of both series, causing contradictory radiocarbon dates. For higher absolute ages, a similar restriction does not exist. But a shift to higher absolute ages by more than about one century requires an exceptionally large decrease of the atmospheric ^{14}C levels, exceeding the observed range of Holocene ^{14}C oscillations.

According to this absolute age placement, the pine series begins at a minimum age of 11,370 dendro-years BP. The pattern of atmospheric ^{14}C variations to this date can be reconstructed (Fig. 1a). The data indicate an Early Holocene ^{14}C level 70 to 90% above the present level. This range of the ^{14}C level does

not agree with values derived indirectly from the absolute age determination of the end of the Younger Dryas in Greenland ice cores⁷. Our results are closer to the ^{14}C values obtained on North American varves⁸ and to results from U/Th dated corals⁹.

In the early Holocene the long-term ^{14}C trend is superimposed by strong century-type variations. This leads to intervals of apparent constant ^{14}C ages for periods of up to 450 dendro-years in the calibration curve (Fig. 2c). These ^{14}C -age plateaux at 10,000, 9,600, 8,750 and 8,200 ^{14}C yr BP cause severe problems in ^{14}C dating for the 10,000 to 9,000 ^{14}C yr BP interval. For example, the 400 ^{14}C yr interval framed by the two plateaux at 10,000 and 9,600 ^{14}C yr BP can represent a minimum of 90 and a maximum of 860 dendro-years.

As a consequence of the plateau at 10,000 ^{14}C yr BP, the conventional radiometric assignment of the transition from Late Glacial to Holocene cannot be defined precisely in absolute terms. The pine tree-ring pattern of this period also does not show a significant change in the long-term trend of tree-ring growth rates to be interpreted as a response to the global climatic change that terminated the Younger Dryas cold phase. Fortunately, a remarkable change of climate is seen in the stable isotope record of wood samples from the German pine chronology.

Five-year segments of early and late wood samples of 15 pine trunks from the Danube river were prepared for ^{13}C and ^2H analysis. The ^2H and ^{13}C technique involves the preparation of pure cellulose nitrate and uses only the isotopically nonexchangeable C-H hydrogen of cellulose¹⁰. All measurements of $^2\text{H}/^1\text{H}$ and $^{13}\text{C}/^{12}\text{C}$ ratios are reported in the conventional δ values, where $\delta = (R_{\text{sample}}/R_{\text{standard}} - 1) \times 1,000$ (‰). The sample series

cover the tree rings 180 to 1,490 of the chronology (Fig. 2b). The long-term isotope trends are confirmed by stepwise overlaps between at least two trees, with the exception of a gap of data between rings 1,145 to 1,385. In cases where the wood structure was partly decayed, 10-yr segments of total wood (early plus late wood) were used. A study of separately prepared early- and late-wood Preboreal pine samples showed neither significant offsets nor an increase in variance compared with the results of the 5- and 10-yr samples. Thus, 10-yr total wood ^{13}C and ^2H averages are given. To minimize possible juvenile effects of the trees affecting the ^{13}C content of the cellulose, the innermost 5 decades of each cross-section were excluded.

Although tree-ring $^{13}\text{C}/^{12}\text{C}$ ratios can vary, even within a single year, some recent studies indicate a significant correlation between $^{13}\text{C}/^{12}\text{C}$ variations of tree rings and climate^{11,12}. The theoretical explanation for drought-induced $\delta^{13}\text{C}$ variations of conifer chronologies in the southwestern United States is the increased stomatal closure during drought periods¹³. This leads to elevated $^{13}\text{C}/^{12}\text{C}$ ratios because the reduction of available CO_2 diminishes the photosynthetic discrimination against ^{13}C in favour of ^{12}C . We interpret the Preboreal pine isotope curve in a similar way as indication of change of climate.

As shown in Fig. 2a, low $\delta^{13}\text{C}$ values of $-24 \pm 0.3\%$ occur at the beginning of the pine series (the first 250 measured dendro-years). We assume this part of the chronology to correspond to the latest phase of Younger Dryas. Coincident with the end of the 10,000- ^{14}C -yr age plateau, the $\delta^{13}\text{C}$ content of the pine wood increases substantially. During the first two centuries of the Preboreal period, the ^{13}C curve shifts to values of about -22.5% . This significant 1.5‰ increase reflects, in our opinion, the global climatic improvement which terminated the Younger Dryas cold phase. This phase thus ended at a minimum age of 10,970 dendro-years BP.

During the subsequent Preboreal period (covered by samples from dendro-year 600 to 1,550 of the pine sequence), $\delta^{13}\text{C}$ remains nearly constant except for a decrease between rings 1,060 and 1,120. Here, $\delta^{13}\text{C}$ drops to another minimum of $\delta^{13}\text{C}$ values of -23.5% .

Although the early Preboreal increase of the $\delta^2\text{H}$ pine tree-ring record lasts longer than that of $\delta^{13}\text{C}$, the general trend of both curves is very similar. The $\delta^2\text{H}$ values are low during the Younger Dryas (-100 to -90%) but increase along with ^{13}C . They continue to rise even after $\delta^{13}\text{C}$ has reached a plateau. At the end of the 1,500-yr tree ring isotope series, a maximum $\delta^2\text{H}$ value of -70% is reached. The late Preboreal $\delta^{13}\text{C}$ decrease is found in the $\delta^2\text{H}$ record as well. The $\delta^2\text{H}$ curve rises again and reaches maximum values of -70% at the end of the Preboreal.

Similar to the ^{13}C variations observed in modern pines, the $\delta^{13}\text{C}$ increase observed in German subfossil pines at the early Holocene is considered to be a consequence of stomatal conductance in the pine needles resulting from higher temperature and/or reduced water supply. With the (isotopic) resistor chain model^{14,15} and reasonable assumptions for the maximum assimilation flux, for the viscous boundary layer, the stomata dimensions, and the isotope enrichment¹² of the fixation reaction, a reduction to about 60% of the maximum stomata aperture explains the observed increase of 1.5‰ in $\delta^{13}\text{C}$ (K. O. Münnich, personal communication).

In terms of atmospheric humidity, application of the humidity- $\delta^{13}\text{C}$ gradient as demonstrated for North American trees¹⁶ would imply a 15% reduction in relative humidity when wood $\delta^{13}\text{C}$ increases by 1‰. The 1.5‰ $\delta^{13}\text{C}$ increase associated with the climatic amelioration following the Younger Dryas phase (YD transition) in southern Germany is of similar magnitude as the 0.8 to 1.8‰ $\delta^{13}\text{C}$ change found between wood samples from glacial origin and modern wood in North America^{16,17}.

The causes for the observed tree $\delta^{13}\text{C}$ change should include (1) global change in $\delta^{13}\text{C}$ values of the atmospheric CO_2 used for photosynthesis, and (2) changes in the degree of isotope fractionation related to the biological response of the tree to

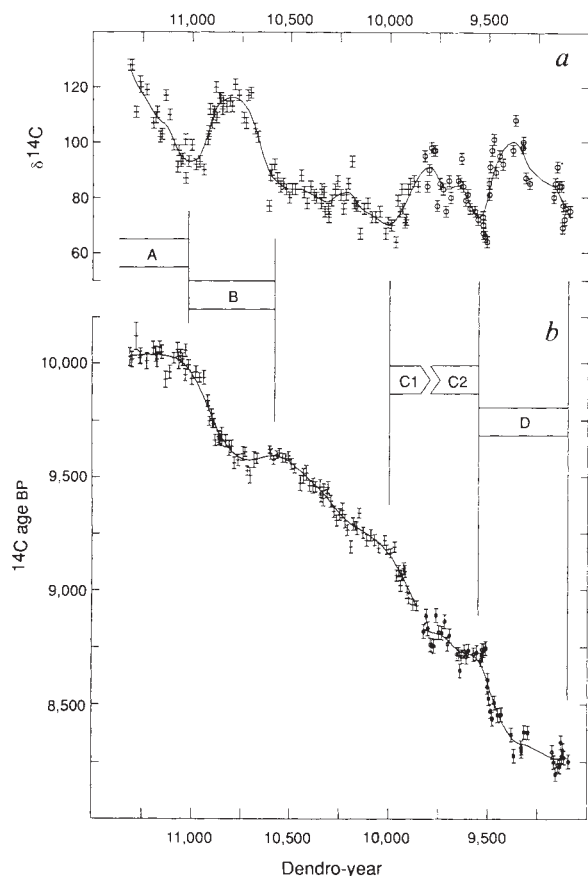


FIG. 1 Atmospheric $\delta^{14}\text{C}$ level (a) and ^{14}C -age calibration curve (b). The four strong ^{14}C oscillations are marked by letters A, B, C and D. The derivation of the dendro timescale includes radiometrical crossmatching of the pine series to the absolute oak at the oscillation C1/C2 (see text).

local climatic change (for example, lower humidity at increasing temperature). Atmospheric $\delta^{13}\text{C}$ is related to the average temperature and $\delta^{13}\text{C}$ value of the surface-water bicarbonate of the world oceans. Global increase in $\delta^{13}\text{C}$ of the surface oceans for the glacial/interglacial transition is in the 0.2 to 0.5‰ range^{18,19}, whereas another 0.4‰ atmospheric $\delta^{13}\text{C}$ increase related to chemical fractionation would be induced by a global warming of 5°C. Thus the full glacial/interglacial atmospheric $\delta^{13}\text{C}$ change is probably less than 1‰, implying only a fractional contribution of perhaps 0.3 or 0.4‰ for the more limited YD transition. Evidently the biological fractionation change dominates the $\delta^{13}\text{C}$ signal.

The palaeoclimatic significance of deuterium in tree ring cellulose is not yet well understood. The synchronous pattern of early Preboreal $\delta^{13}\text{C}$ and $\delta^2\text{H}$ pine series indicates that both isotope signals of subfossil pines reflect a global improvement of the climate. Furthermore, the timing of the $\delta^2\text{H}$ change in the pine wood at the YD transition coincides with the $\delta^{18}\text{O}$ increase in marine sediments²⁰ and with climate-induced changes in Swiss and southern German pollen diagrams²¹.

Finally, the observed variations in deuterium agree remarkably well with the change in deuterium contents of Late Glacial to modern groundwaters in this area²². A principal cause for the 25‰ tree $\delta^2\text{H}$ increase may well be a change in the isotope composition of the local precipitation. In which case, the $\delta^{18}\text{O}$ change would be one eighth of the $\delta^2\text{H}$ change, or 3‰, in good

agreement with the 2.3‰ and 3.5‰ $\delta^{18}\text{O}$ increases encountered for the YD transition in the Swiss lakes Tourbiere de Chirens and Gerzensee^{23,24}, respectively.

Our independent time series of tree ring growth, $\delta^{13}\text{C}$, $\delta^2\text{H}$ and ^{14}C establishes a precise timescale for the start of the climatic improvement at the beginning of the Holocene. We suggest that the synchronous and significant increase in $\delta^{13}\text{C}$ and $\delta^2\text{H}$ of the German pine tree ring record indicates the late Glacial Holocene boundary. The $\delta^{13}\text{C}$ and $\delta^2\text{H}$ transitions lasted 100 to 200 years, and 450 years, respectively. According to our cross-match of high-precision ^{14}C ages of the pine and the absolute oak chronology, the Late Glacial must have ended at a minimum age of 10,970 dendro-years BP.

Received 11 April; accepted 20 August 1991.

1. Becker, B. *et al.* *Antiqua* **11**, 68 (1985).
2. Pilcher, J. R., Baillie, M. G. L., Schmidt, B. & Becker, B. *Nature* **312**, 150–152 (1984).
3. Becker, B. & Schmidt, B. *PACT* **29**, (in the press).
4. Stuiver, M. & Kra, R. (eds) *Radiocarbon* **28**, 28, (1986).
5. Stuiver, M., Brazunias, Th. F., Becker, B. & Kromer, B. *Quat. Res.* **35**, 1–24 (1991).
6. Stuiver, M. & Brazunias, Th. F. *Nature* **338**, 405–408 (1989).
7. Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284–291 (1986).
8. Stuiver, M. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 57–70 (Yale University Press, New Haven, 1971).
9. Bard, E., Hamelin, B., Fairbanks, R. G. & Zindler, A. *Nature* **345**, 405–410 (1990).
10. Epstein, S., Yapp, C. J. & Hall, J. H. *Earth planet. Sci. Lett.* **30**, 241–251 (1976).
11. Leavitt, St W. & Long, A. *Global biochem. Cycles* **2**, 189–198 (1988).
12. Francey, R. J. & Farquhar, G. D. *Nature* **297**, 28–31 (1982).
13. Leavitt, St W. & Long, A. *Water Resource Bull.* **25**, 341–347 (1989).
14. Vogel, J. C. *Sitzungsber. der Heidelberger Akademie der Wissenschaften, Math.-Nat. Klasse*, Jg. 1980, 3. Abh. (Springer, Berlin, 1980).
15. Münnich, K. O. *et al.* in *Turbulent Fluxes through the Sea Surface, Wave Dynamics and Prediction* (eds Favre, A. & Hasselmann, K.) (Plenum, New York, 1978).
16. Stuiver, M. & Brazunias, Th. F. *Nature* **327**, 58–60 (1987).
17. Yapp, C. J. & Epstein, S. *Earth planet. Sci. Lett.* **34**, 333–350 (1977).
18. Shackleton, N. J. & Pisias, N. G. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaea to Present* (eds Sundquist, E. T. & Broecker, W. S.) Geophys. Monograph **32**, 303–317 (1985).
19. Curry, W. B. & Crowley, T. J. *Palaeoceanography* **2**, 489–517 (1987).
20. Bard, E. *et al.* *Nature* **328**, 791–794 (1987).
21. Lotter, A. *Diss. Botanicae* **124** (Cramer, Berlin-Stuttgart, 1988).
22. Moser, H. *et al.* *Palaeoclimates and Palaeowaters: A Collection of Environmental Isotope Studies* (IAEA, Vienna, 1983).
23. Eicher, U. & Siegenthaler, U. *Boreas* **5**, 109–117 (1976).
24. Eicher, U., Siegenthaler, U. & Wegmüller, S. *Quat. Res.* **15**, 160–170 (1981).

ACKNOWLEDGEMENTS. We thank K. O. Münnich and M. Stuiver for discussion. This work was funded by the National Climate Program supported by the German Federal Ministry of Research and Technology.

Isotope evidence for the involvement of recycled sediments in diamond formation

C. S. Eldridge^{*†}, W. Compston[†], I. S. Williams[†],
J. W. Harris[‡] & J. W. Bristow[§]

^{*} Geology Department, Australian National University, GPO Box 4, Canberra, Australian Capital Territory 2601, Australia

[†] Research School of Earth Sciences, Australian National University, GPO Box 4, Canberra, Australian Capital Territory 2601, Australia

[‡] Department of Geology and Applied Geology, University of Glasgow, Glasgow G12 8QQ, UK

[§] Research Laboratory, Anglo American Corporation Ltd, PO Box 106, Crown Mines 2025, Republic of South Africa

DIAMONDS, as chemically robust containers of material from the mantle, have provided important information regarding the evolution of the Earth's interior¹; yet despite much effort^{2–5}, the origins of diamonds themselves have remained controversial^{10–13}. Sulphur isotope ratios measured in most mantle or meteoritic samples exhibit minimal deviation from the meteoritic standard, whereas the ratios in crustal materials deviate by as much as $\pm 70\%$ (ref. 14); for this reason, isotopic studies of sulphide mineral inclusions in diamonds and mantle material are an attractive means of investigating the possible role of crust–mantle interaction in

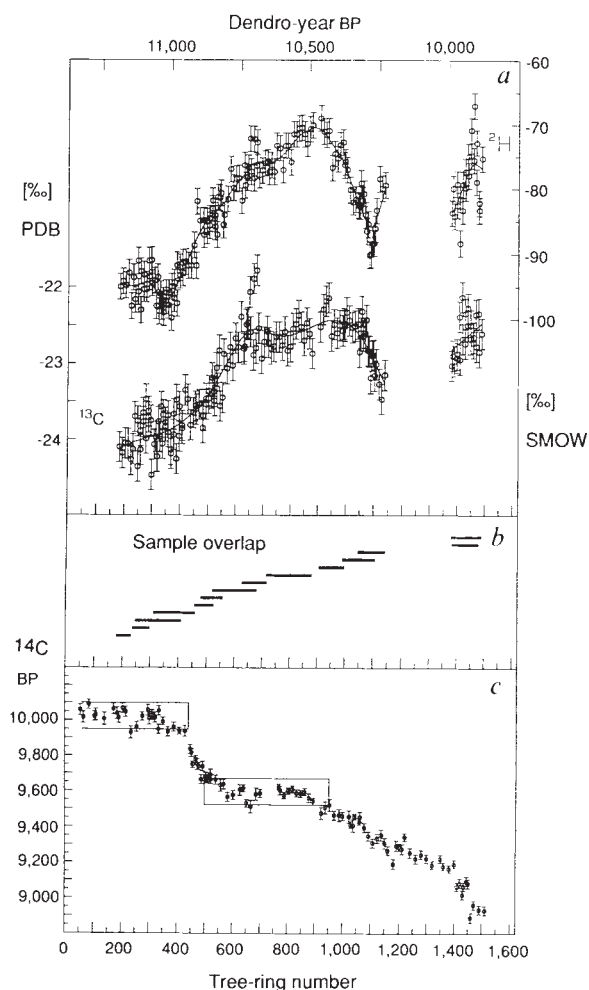


FIG. 2 *a*, $\delta^2\text{H}$ and $\delta^{13}\text{C}$ tree-ring records at the transition from Younger Dryas to Preboreal. *b*, Replication of the isotope series by stepwise overlapping of individual trees. *c*, ^{14}C -age plateaux in the Preboreal pine series at 10,000 BP and 9,600 BP.

the formation of diamonds¹⁵. Recent ion microprobe studies have found departures from mantle values in the sulphur isotope compositions of individual diamond inclusions¹⁶ and some mantle rocks¹⁷, which would be consistent with subduction of altered ocean crust into the region of diamond growth. Here we present new sulphur and lead isotope data from diamond sulphide inclusions, which suggest that sedimentary material might also be subducted into the mantle and that the crustal recycling process has been operative for at least 1,000 million years.

Sulphide inclusions in diamonds are rarely more than 100 μm across, yet many contain one or all of the minerals pyrrhotite, pentlandite, chalcopyrite and pyrite. Pyrrhotite and pentlandite dominate the inclusions, with chalcopyrite (with or without minor pyrite) occurring mainly as rims. Inclusions with both pyrrhotite and pentlandite often have exsolution textures where areas of pentlandite may vary from $<1\ \mu\text{m}$ to $\sim 35\ \mu\text{m}$ across. The intergrowths are often too closely spaced to analyse independently even with the electron microprobe, and the mixture is simply referred to as monosulphide solid solution (m.s.s., for example in Table 1).

Although analytical techniques for sulphur isotope measurements described earlier^{18,19} could be used for pure phases, to analyse m.s.s. variable instrumental corrections had to be applied if the craters produced by the ion beam simultaneously sampled pyrrhotite and pentlandite. The fractionation of ^{32}S and ^{34}S during sputtering of pyrrhotite and pentlandite differs by $\sim 1\%$ and, therefore, each analytical crater created by the ion microprobe was analysed for Fe, Cu, Ni and S contents by electron microprobe to allow proper correction for sputtering fractionation. In this study, sulphur isotope ratios were found to be reproducible within analytical uncertainties on all of these matrices, and individual sulphide inclusions appear isotopically homogeneous, with even multiphase inclusions yielding single isotopic values in all phases (Table 1). Sulphur isotope values had a spread of 25% (-11 to 14% , Table 1) which is the largest reported for mantle samples. The overall mean is 0.4% , however, and such heterogeneity could simply have gone unnoticed in a conventional isotope study. Half of the kimberlite pipes, Koffiefontein, Roberts Victor, Finsch and Premier, yielded sulphides

with relatively limited ranges of sulphur isotope compositions near 0% (Table 1). The other half yielded groups of inclusions with a much greater range of compositions: Sierra Leone (8 – 14%), Mwadui (-2 to 10%), Orapa (-11 to 2%) and Jagersfontein (-10 to 7% , Table 1; Fig. 1). It is well known that many diamond pipes contain materials of peridotitic and eclogitic affinities, but these sulphides are from diamonds without silicate inclusions so their paragenesis cannot be determined in the traditional manner. It has been suggested that sulphide paragenesis is reflected in the nickel content of the inclusion²⁰. Comparison of sulphur isotope data and cation composition of the inclusions showed that $\delta^{34}\text{S}$ values differ significantly from zero (by more than $\pm 4\%$) only in inclusions with nickel contents of less than $8\ \text{wt}\%$ (Fig. 2), which is a proposed maximum nickel content for sulphides of eclogitic paragenesis²⁰. Samples with more than 8% , peridotitic sulphides, show little deviation from 0% .

The lead contents of the inclusions range from <1 to $1,500\ \text{p.p.m.}$, according to lead count rates relative to galena. Those inclusions with more than $5\ \text{p.p.m.}$ lead (17 of the 33 inclusions, Table 1) were analysed for lead isotope composition using a procedure slightly modified from that used in zircon dating^{21,22}. Backgrounds were determined $0.1\ \text{a.m.u.}$ above the mass of each lead isotope to monitor scattered ions. Broken Hill galena was used as a lead standard and zircon SL3 as a standard for monitoring approximate uranium and thorium contents. Surface lead contamination was minimized by ion-cleaning each analytical area immediately before analysis. The fractionation of lead isotopes is much less matrix-dependent than that of sulphur isotopes²³, so no corrections were applied for pentlandite relative to pyrrhotite (for example, P205B, Table 1). Measured ratios for coexisting sulphides were reproducible within analytical error.

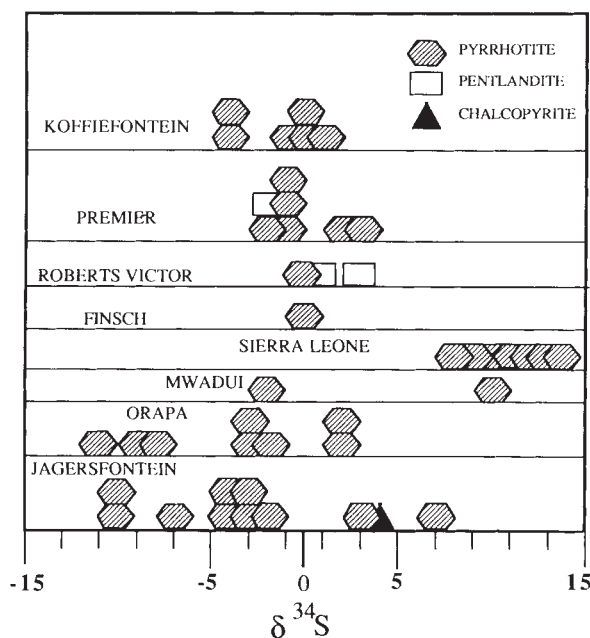


FIG. 1 Sulphur isotope compositions, plotted in parts per thousand relative to the Canyon Diablo troilite standard, determined for the inclusions from eight kimberlite pipes located in three cratons of the African continent. For simplicity, here and in Fig. 2, the symbol for pyrrhotite encompasses cation compositions that fall in the monosulphide solid solution field as well.

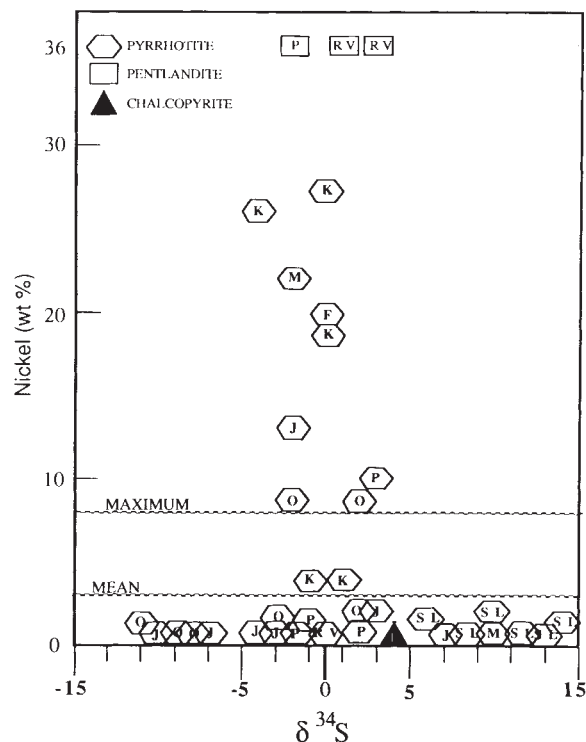


FIG. 2 Sulphur isotope compositions, as in Fig. 1, plotted against weight per cent nickel for the same ion-probe crater. The letters represent: K for Koffiefontein, P for Premier, RV for Roberts Victor, F for Finsch, SL for Sierra Leone, M for Mwadui, O for Orapa and J for Jagersfontein. The mean and maximum lines are from Sobolev²⁰ and represent the relevant values of nickel contents for sulphide inclusions from diamonds of probable eclogitic affinity. The significant sulphur isotope variability is within the eclogitic compositional range.

TABLE 1 Ion microprobe sulphur and lead isotope data from diamond inclusion sulphides

Sample name	Ni wt%	$\delta^{34}\text{S}$	$^{208}\text{Pb}/^{206}\text{Pb}$	Error	$^{207}\text{Pb}/^{206}\text{Pb}$	Error	$^{204}\text{Pb}/^{206}\text{Pb}$	Error	p.p.m. Pb
K 202 (po)	4.0	-1							
(ccp)	0		1.770	0.233	0.576	0.127	0.127	0.058	8
K 205 (ccp)	1.2	-4	2.457	0.264	0.972	0.134	0.087	0.036	9
K 209 (m.s.s.)	18	0							
K 210 (m.s.s.)	25.8	-4	2.254	0.137	0.845	0.062	0.037	0.024	37
			2.545	0.122	0.947	0.053	0.060	0.021	59
K 211 (m.s.s.)	3.6	1							
K 226 (m.s.s.)	27.7	0	2.204	0.134	0.841	0.063	0.060	0.023	35
P 205A (po)	0.4	-1							
(pnt)	36.0		1.308	0.274	0.442	0.125	<0.153	0.153	7
	(3%)								
P 205B (po)	0.4	-2	2.505	0.342	0.891	0.157	0.121	0.057	5
(pnt)	36.0		2.261	0.188	0.838	0.085	0.077	0.034	23
	(5%)								
P 216 (m.s.s.)	9.9	3							
P 217 (po)	1.0	-1							
(po)	1.0	2							
(pnt)	37.4	-2							
	(4%)								
P 218 (po)	0.2	-1							
RV 72C (po)	0.4	0	1.950	0.062	0.760	0.029	0.041	0.004	98
RV 201 (pnt)	35.8	3	2.455	0.068	0.981	0.032	0.068	0.010	147
RV 202 (pnt)	34.3	1	2.176	0.018	0.930	0.009	0.062	0.002	1,426
			2.158	0.017	0.938	0.009	0.065	0.002	1,499
F 101 (m.s.s.)	20.3	0	2.420	0.056	1.053	0.028	0.078	0.007	183
			2.434	0.049	1.044	0.025	0.076	0.006	248
SL 1 (po)	0.3	11							
	0.3	9							
SL 3 (po)	0.4	12							
	0.4	13							
SL 21 (po)	1.1	8	1.771	0.186	0.498	0.100	0.026	0.031	11
SL 26 (po)	1.6	14	2.042	0.062	0.844	0.032	0.058	0.007	108
	1.6	10	2.104	0.096	0.772	0.049	0.036	0.014	53
M 1B (m.s.s.)*	21.8	-2							
M 8 (po)	0.4	10							
O 209A (po)*	1.1	-9	1.236	0.228	0.170	0.137	0.010	0.023	5
	1.1	-8							
O 209B (po)*	1.6	-11							
O 211 (po)	1.4	-3	2.201	0.087	0.877	0.044	0.063	0.011	64
			2.207	0.074	0.934	0.039	0.059	0.008	89
O 212 (po)	1.4	-3	2.018	0.096	0.880	0.051	0.053	0.012	45
			1.977	0.103	0.873	0.057	0.054	0.013	39
O 214 (m.s.s.)	9.1	-2	2.119	0.093	0.887	0.047	0.030	0.012	54
O 214	9.2	2	2.428	0.102	1.002	0.051	0.054	0.015	65
O 217 (m.s.s.)	2.2	2							
J 209 (po)	0.2	-10							
	0.2	-10							
J 215 (po)	2.4	3							
	2.0	7							
J 215 (ccp)	0.1	4	1.812	0.078	0.670	0.037	0.056	0.010	55
			1.922	0.051	0.708	0.024	0.043	0.006	150
J 217 (po)*	0.3	-7							
	0.3	-4							
	0.3	-3							
J 218 (m.s.s.)	12.9	-2	2.334	0.079	0.981	0.041	0.037	0.011	90
			2.384	0.067	1.011	0.034	0.056	0.010	126
J 219 (po)	0.2	-4							
	0.2	-3							

K, Koffiefontein; P, Premier; RV, Roberts Victor; F, Finsch; SL, Sierra Leone; M, Mwadui; O, Orapa; J, Jagersfontein. K, P, RV, F and O are from the Kaapvaal craton; M is from the Tanzania craton; and SL is from the West Africa craton. po, pyrrhotite; ccp, chalcopryrite; pnt, pentlandite; m.s.s., monosulphide solid solution. Weight per cent nickel (Ni wt%) analysed by electron microprobe in the ion microprobe crater; * indicates a broad-beam analysis. Values in parentheses are average nickel content estimate for total inclusion of coarse-grained pentlandite plus pyrrhotite. $\delta^{34}\text{S}$, the sulphur isotopic composition (‰) relative to the Canyon Diablo troilite standard. Where duplicate data are listed, these come from two spots on the same inclusion. Values are rounded to the nearest whole number, as errors are $\pm 2\%$ (2σ). Errors in the lead isotope ratios are 1σ . Uranium contents, to the extent that they could be estimated from measured count rates, ranged from 0 to 8 p.p.b. with a mean of 2 p.p.b. Thorium abundances were similarly low, varying from 2 to 22 p.p.b. with a mean of 15 p.p.b. In 100 Myr, decay of these low levels of parent isotopes would essentially leave the lead ratios unaltered and even in 1 Gyr they would contribute less than 1 p.p.b. of any one isotope, so the lead isotope ratios were left uncorrected for *in situ* decay. The measured levels of lead, uranium and thorium are similar to those found elsewhere³.

The evolution of terrestrial common lead through time, for average U/Pb and Th/Pb ratios²⁴, is illustrated in Fig. 3. Simple mantle leads should lie on this curve or in the area below the curve, with their lower limits determined by the geochron, a straight line connecting the modern day (0 age) with the time of formation of the planet (4.5 Gyr). Roughly half of the lead isotopic compositions found in the diamond sulphide inclusions fall on the lead evolution curve or within the area bounded by the curve and the geochron. Those which lie off the curve towards the origin do not meet the requirements of simple mantle leads, and it is interesting to note that these inclusions are also the ones with low nickel contents (eclogitic) and host to the sulphur isotope compositions that differ greatly from 0‰ (Figs 1 and 3, Table 1).

The sulphur and lead data, together with the nickel contents of the sulphide inclusions, point to an internally consistent set of origins for the above two groups. All three of the geochemical characteristics of group 1 or peridotitic material ($\delta^{34}\text{S}$ values $\sim 0\%$, lead isotope ratios between the evolution curve and the geochron, and Ni contents $>8\%$) could have resulted from formation in an uncontaminated mantle. Group 2 or eclogitic material ($\delta^{34}\text{S}$ values $<-4\%$ or $>4\%$, lead isotope ratios lying off the evolution curve toward the origin, and Ni $<8\%$) reflects a crustal setting, where $\delta^{34}\text{S}$ values significantly below zero are consistent with a sedimentary origin, perhaps resulting from biogenic sulphate reduction, or may even record inorganic processes such as hydrothermal precipitation in a relatively low-temperature, oxidizing environment. Those well above 0‰ might also have been formed in a sedimentary setting as a result of biogenic effects, or may represent hydrothermally altered sea-floor basalt²⁵. The anomalous lead data suggest a source with higher than average U/Pb ratio (μ value). If the age of the source is assumed to be 1 Gyr, based on the cluster of eclogitic leads at the point on the lead evolution curve (Fig. 3), and the most discordant lead isotope composition (Fig. 3; O209A, Table 1) is used as the end point, the appropriate ratio is $\mu \approx 300$. Such a high μ value would not be expected in the mantle, but

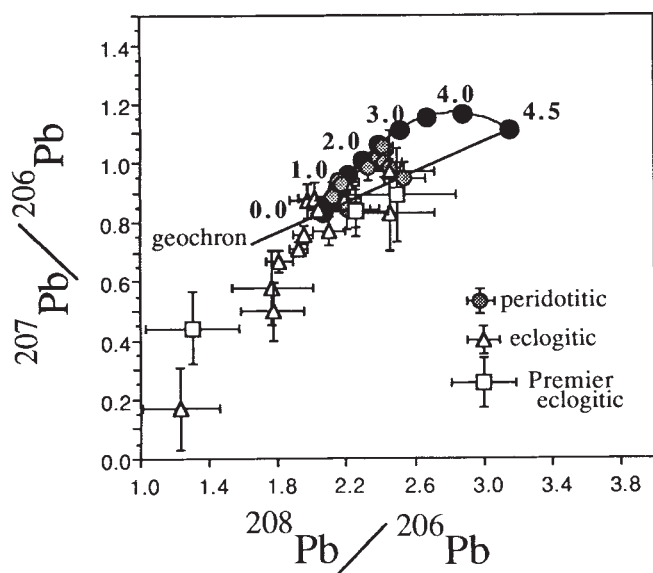


FIG. 3 Using the eclogitic against peridotitic classification from Fig. 2, lead isotope compositions have been plotted in $^{208}\text{Pb}/^{206}\text{Pb}$ against $^{207}\text{Pb}/^{206}\text{Pb}$ space with respect to the lead evolution curve²⁴. Model ages are shown in heavy numerals. Many of the peridotitic sulphides fall on the curve, grown from perhaps 2 Gyr ago with average crystal μ . Those falling below the curve may have formed in environments of lower μ . In either case, the compositions do not fall out of the realm of possible normal mantle source regions. Eclogitic inclusions, in some cases, also show compositions which fit on, or near to, the lead evolution curve. In contrast to the peridotitic inclusions, however, many lie off the curve towards the origin. These require growth in a high μ environment, possibly evolved crust.

is very likely to be crustal and, as is consistent with the implications of the sulphur isotope ratios, of a sedimentary origin. Finally, it is fitting that the eclogitic or crustal material should have a lower nickel content than the peridotites or mantle samples.

The lead data can also be used to give some indication of the age of the material incorporated into diamonds, and the age of the inclusions need not necessarily match the known age of emplacement of the kimberlite pipe. Eruption ages for seven of the kimberlites studied are ~ 100 Myr, with Premier being considerably older at 1.2 Gyr. Lead isotope compositions of some of the peridotitic samples indicate, however, that there is some very old lead in the inclusions, reaching $\sim 2\text{--}2.5$ Gyr (Fig. 3). This may represent only a maximum age for these particular diamonds, as old lead could simply have been incorporated by the sulphide during growth. The lead in many of the eclogitic inclusions seems to be fairly young, although it is likely that these radiogenic compositions reflect mixing between common and radiogenic lead, rather than directly reflecting an age of formation. The inclusions with the more radiogenic compositions are also the ones with lower lead contents (5–50 p.p.m., Table 1) whereas the less radiogenic lead isotope compositions occur in inclusions with roughly an order of magnitude more lead (35–1,500 p.p.m., Table 1). Naturally, the inclusions with the lower lead contents would be more likely to reflect contamination with anomalous lead, should mixing have occurred.

Our results support the findings of earlier work and, furthermore, establish a logical link between systematic trends in stable and radiogenic isotope compositions of mantle material, removing some of the uncertainty about the role of plate tectonics in diamond formation. It was earlier pointed out^{16,17} that the sulphur isotope composition of the mantle is heterogeneous and that input of oceanic crust might account for such behaviour. It was also suggested from very negative carbon isotopic compositions^{1,26,27} and $f\text{O}_2$ modelling⁹ that sediments were placed into the mantle. Our results demonstrate that the sulphur isotope composition of the mantle is even more diverse than previously thought (-11 to 14%) and that the range of variability easily encompasses addition of oceanic crustal and sedimentary components. Lead in diamond inclusions has been shown to have accumulated in low- μ environments³, and it now seems that some lead currently in the mantle may have come from unexpectedly high- μ regions (~ 300). This idea is compatible with a crustal origin, possibly continental crust. It may also be more common than expected^{3,4,6} for diamonds to contain material older than their eruption age, with both eclogitic and peridotitic diamonds being xenocrysts. It is clear that recycled crust has been returned to the surface for more than 1,000 million years and in at least three cratons of one of the world's most diamond-rich continents, as is consistent with previous suggestions linking subduction to diamond genesis²⁸. By correlating sulphur, lead and nickel geochemistries, it may now be possible to establish the relationship between the diamond content of kimberlites and the nature of their source material. □

Received 14 May; accepted 23 August 1991.

- Kesson, S. E. & Ringwood, A. E. *Chem. Geol.* **78**, 97–118 (1989).
- MacGregor, I. D. & Manton, W. I. *J. geophys. Res.* **91**, 14063–14079 (1986).
- Kramers, J. D. *Earth planet. Sci. Lett.* **42**, 58–70 (1979).
- Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198–202 (1984).
- Richardson, S. H., Erlank, A. J. & Hart, S. R. *Earth planet. Sci. Lett.* **75**, 116–128 (1985).
- Richardson, S. H., Erlank, A. J., Harris, J. W. & Hart, S. R. *Nature* **346**, 54–56 (1990).
- Burgess, R., Turner, G., Laurenzi, M. & Harris, J. W. *Earth planet. Sci. Lett.* **94**, 22–28 (1989).
- Turner, G., Burgess, R. & Bannon, M. *Nature* **344**, 653–655 (1990).
- Bailhaus, C., Berry, R. F. & Green, D. H. *Nature* **348**, 437–440 (1990).
- Deines, P., Harris, J. W. & Gurney, J. J. *Geochim. cosmochim. Acta* **51**, 1227–1243 (1987).
- Deines, P. *Geochim. cosmochim. Acta* **44**, 943–961 (1980).
- Deines, P. & Wickman, F. E. *Geochim. cosmochim. Acta* **39**, 547–557 (1975).
- Deines, P., Harris, J. W., Robinson, D. N., Gurney, J. J. & Shee, S. R. *Geochim. cosmochim. Acta* **55**, 515–524 (1990).
- Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* 2nd Edn. (ed. Barnes, H. L.) 509–567 (Wiley, New York, 1979).
- Tsai, H., Shieh, Y. & Meyer, H. O. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* Vol. 2 (eds Boyd, F. R. & Meyer, H. O. A.) 87–103 (American Geophysical Union, Washington, DC, 1979).

16. Chaussidon, M., Albarède, F. & Sheppard, S. M. F. *Nature* **330**, 242–244 (1987).
17. Chaussidon, M., Albarède, F. & Sheppard, S. M. F. *Earth planet. Sci. Lett.* **92**, 144–156 (1989).
18. Eldridge, C. S., Compston, W., Williams, I. S. & Walshe, J. L. & Both, R. A. *Int. J. Mass Spectrom. Ion Proc.* **54**, 40–59 (1987).
19. Eldridge, C. S., Compston, W., Williams, I. S. & Walshe, J. L. *US Geol. Surv. Bull.* **1890**, 163–173 (1989).
20. Sobolev, N. V. in *Kimberlite Occurrence and Origin: A Basis for Conceptual Models in Exploration Geology* Dept and University Extension Univ. of Western Australia Publ. No. 8 (eds Glover, J. E. & Harris, P. G.) 275–287 (1984).
21. Compston, W., Williams, I. S. & Meyer, C. *J. geophys. Res.* **89**, B525–B534 (1984).
22. Williams, I. S. & Claesson, S. *Contrib. Mineral. Petrol.* **9**, 205–217 (1987).
23. Shimizu, N. & Hart, S. R. *J. appl. Phys.* **53**, 1303–1311 (1982).
24. Cumming, G. L. & Richards, J. R. *Earth planet. Sci. Lett.* **28**, 155–171 (1975).
25. Sakai, H., Des Marais, D. J., Ueda, A. & Moore, J. G. *Geochim. cosmochim. Acta* **48**, 2433–2441 (1984).
26. Sobolev, N. V. in *Kimberlite Occurrence and Origin: A Basis for Conceptual Models in Exploration Geology* Dept and University Extension Univ. of Western Australia Publ. No. 8 (eds Glover, J. E. & Harris, P. G.) 213–226 (1984).
27. Jaques, A. L. et al. in *Kimberlites and Related Rocks, 2, Their Mantle/Crust Setting, Diamonds, and Diamond Exploration*, Proc. 4th Int. Kimberlite Conf., 966–989 (Blackwell, London, 1989).
28. Helmstaedt, H. & Gurney, J. J. in *Kimberlites and Related Rocks, 1, Proc. 3rd Int. Kimberlite Conf.* (ed. Kornprobst, J.) 425–434 (Elsevier, Amsterdam, 1984).

ACKNOWLEDGEMENTS. We thank P. D. Kinny, T. R. Ireland, W. F. McDonough and S. E. Kesson of RSES for discussions and comments on the manuscript. Initial electron microprobe analyses were made by P. Hill and S. Kearns at the University of Edinburgh Camebax facility. The electron microprobe data following ion microprobe analyses were obtained with the assistance of N. G. Ware of RSES. C.S.E. acknowledges support through a Research Fellowship provided by the Australian Research Council and a Visiting Fellowship at RSES.

Oscillations and chaos in the dynamics of a perennial grass

David Tilman & David Wedin

Department of Ecology, Evolution and Behavior, University of Minnesota, Minneapolis, Minnesota 55455, USA

ECOLOGICAL models describe the conditions required for chaotic population dynamics^{1–7}, but there are few studies with which to test these predictions^{4,8–10}, and none for perennial plants. In a five-year experiment, a perennial grass exhibited numerous traits associated with chaotic dynamics. Biomass oscillations were greater on more productive soils, with plots on the richest soils exhibiting a 6,000-fold crash. Curves relating the biomass one year to that in the previous year had the peaks and steep slopes associated with chaos, and the dependence of plant biomass on productivity became fuzzy over time. These dynamics resulted from the time-delayed inhibitory effect of plant litter on subsequent growth. A model incorporating litter inhibition predicts oscillations and chaos as productivity increases.

When ecologists sample populations, the resulting relationships are often blurred. This variance may result from sampling error and environmental variation or from deterministic processes that yield oscillations or chaos^{8–11}. Although it is difficult to distinguish between noise and chaos^{9,10}, chaotic dynamics have several distinguishing features^{8–11}. In models such as the discrete logistic, populations maintain a stable density when their regulatory force is small relative to the time delay with which it acts. At higher intensities of the regulatory force, populations exhibit sustained oscillations. At even higher intensities, these periodic orbits give way to aperiodic dynamics called chaos^{1–3,11}. These dynamics can be described using a graph, called a one-dimensional map, of N_{t+1} (population size at time $t+1$) versus N_t . The slope of this curve at the equilibrium line ($N_t = N_{t+1}$) becomes steeper as the regulatory force increases. When sampling noise is superimposed, the one-dimensional map should be increasingly well defined as chaos is approached⁸. Moreover, bifurcations and chaos mean that a wide range of population sizes become associated with the same intensity of the regulatory force³, thus leading to a fuzzy dependence of population size on the regulatory force⁸.

All of these characteristics were displayed by monocultures of *Agrostis scabra* planted at two different initial densities on 10 different soil mixtures. Populations growing on unproductive,

low-nitrogen soils maintained fairly stable above-ground biomasses, those on richer soil oscillated more, and those on the richest soil exhibited a 6,000-fold crash (Fig. 1). None of the other four species growing in this garden exhibited a crash in 1988, suggesting that it was not caused by environmental conditions.

When data from each of the 10 soil mixtures were analysed separately, graphs of B_{t+1} (biomass at time $t+1$) versus B_t had clear, steeply peaked curves for high nitrogen soils (Fig. 2a, b), but had shallow slopes at the equilibrium line and little or no pattern for low-nitrogen soils (Fig. 2c). For the five lowest-nitrogen soil mixtures, none of the slopes differed significantly from zero, but three were significantly negative ($P < 0.05$) for the five richest mixtures. Observed slopes were significantly negatively correlated with soil nitrogen (Fig. 2c). The two most nitrogen-rich soil mixtures had slopes of -3.1 and -3.6 . By comparison, two different forms of the discrete logistic model become chaotic at slopes more negative than -1.57 or -1.69 (ref. 3) and a model of plant-litter dynamics becomes chaotic at -2 (legend to Fig. 2).

In 1986 (Fig. 3a) and 1987, biomass of *Agrostis* was very linearly dependent on soil total nitrogen (1986: $R^2 = 0.92$, $N = 60$, $P < 0.001$; 1987: $R^2 = 0.84$, with high and low seed densities combined for each year) and was independent of the initial seeding density. This relationship was considerably weaker in 1988 ($R^2 = 0.04$) and 1989 ($R^2 = 0.01$). By 1990, the initially tight linear dependence had been replaced by a wedge of data points ($R^2 = 0.32$, $N = 60$, $P < 0.001$), and plots planted at high and low seed densities had diverged (Fig. 3b). This is consistent with the tendency for bifurcations and chaos to cloud the relationship between population size and the intensity of the regulatory force, and to magnify small initial differences.

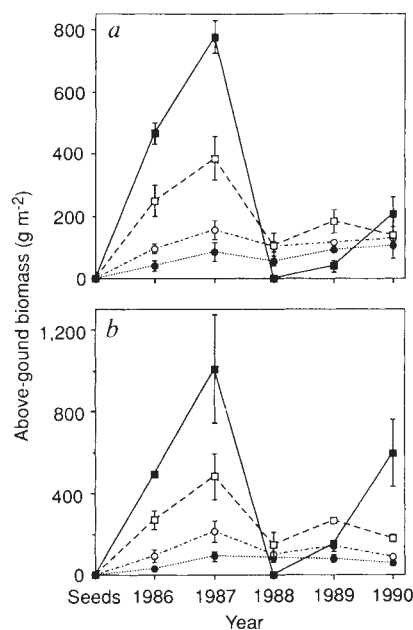


FIG. 1 Above-ground living biomass (mean, s.e.) of *Agrostis* in high seed density monocultures (a) and low seed density monocultures (b). Soil mixtures divided into four groups: N1, three soil mixtures with lowest total soil N (●); N2, three next richer mixtures (○); N3, three next richer mixtures (□); N4, richest mixture (■).

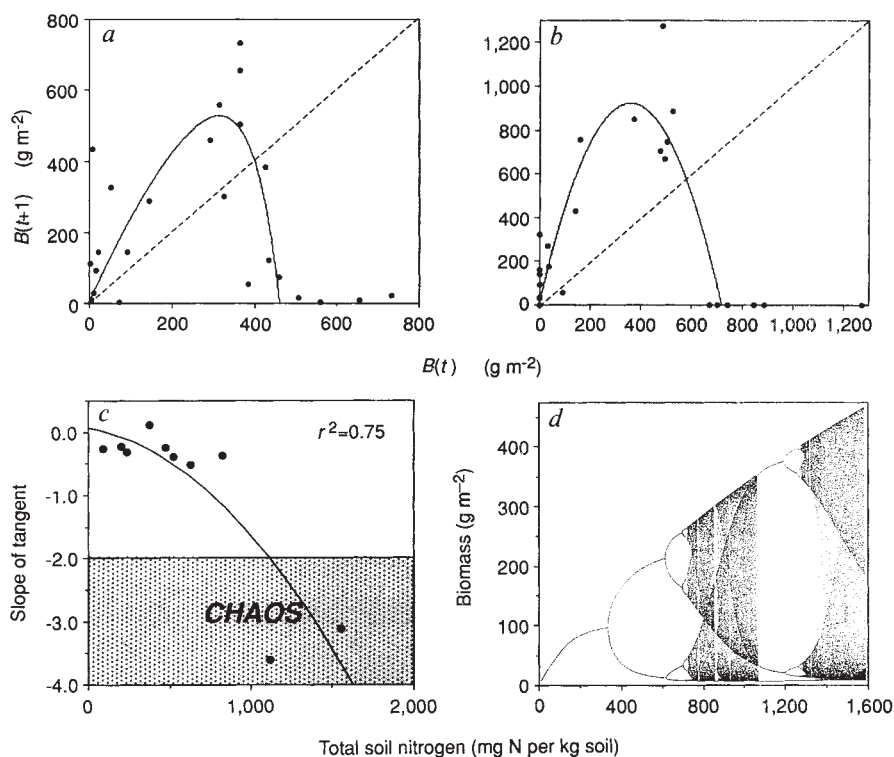
METHODS. *Agrostis* monocultures were seeded in May 1986 to have 3,000 (high seed density) or 600 (low seed density) seedlings m^{-2} , with four high and two low seed density $0.75 m \times 0.75 m$ replicates each in 10 soil mixtures at Cedar Creek Natural History Area, Minnesota, USA. Soils created, watered and fertilized as in ref. 19. Mass of germinated seed in spring 1986 ('Seeds', above) and living biomass in late July each year from samples from different plot subsections each year.

FIG. 2 *a, b*, B_{t+1} (living biomass in year $t+1$) versus B_t for two richest soil mixtures (*a*, 1,116 mg N per kg soil; *b*, 1,550 mg N per kg soil). The 24 data points represent four annual intervals in each of six replicates. Curves were fitted using nonlinear regression ($y = ax - bx^2 - e^{-cx}$, with $y \geq 0$). Equilibrium line, $B(t) = B(t+1)$, is dashed. *c*, Slopes (at equilibrium line) of lines tangent to B_{t+1} versus B_t curves fitted to data from each of 10 soil mixtures. Slopes steeper than -2 give chaos for the model described below. *d*, Bifurcation diagram for a model of plant inhibition by litter. This illustrates the probability of different population biomasses for soils with different total soil nitrogen levels. The model is:

$$B_{t+1} = cN[e^{(a-bL_t)}]/[1 + e^{(a-bL_t)}]$$

$$L_{t+1} = L_t^2/(L_t + d) + ckN[e^{(a-bL_t)}]/[1 + e^{(a-bL_t)}]$$

where B is living biomass, L is litter mass, N is total soil nitrogen, t is time (year), and constants $a, b, c, d > 0$, and $0 < k < 1$. B attains its N-determined equilibrium (cN) in a single year in the absence of litter but is reduced below this by litter. Litter decay is determined by d . Litter production is k times living biomass. Results shown are for $a=5$, $b=0.1$, $c=0.5$, $d=100$ and $k=0.4$, with the model solved at 1,500 different values of N . At each value, results of the first 1,000 years were not used, and those for the next 200 years were graphed against N . This model can be expressed in a simpler, dimensionless form: $X_{t+1} = X_t^2/(1 + X_t) + \alpha/(1 + e^{\beta X_t - \gamma})$ and $Y_{t+1} = \alpha/(1 + e^{\beta X_t - \gamma})$, where X (litter) is L/d , Y (plant biomass) is kB/d , α is ckN/d , β is db , and γ is a . This model illustrates the qualitative effects of a time-delayed litter



feedback on plant dynamics, but is not intended as a description of the dynamics of *Agrostis*.

In total, our results suggest that *Agrostis* biomass underwent oscillations of increasing amplitude, leading to chaos as productivity increased. A likely cause was the inhibition of growth by accumulated litter (the dead plant material produced at the end of a growing season). A dense litter layer mulches the soil surface and intercepts light, thus inhibiting growth until it decays¹²⁻¹⁶. By 1988, litter mass increased significantly with total soil

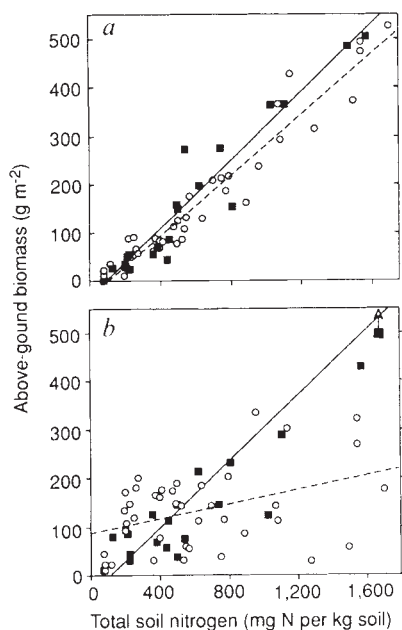


FIG. 3 *a*, Dependence of above-ground living biomass on total soil nitrogen in 1986, the first year of the experiment, for low seed density plots (■) and high seed density plots (○). *b*, Similar results in 1990.

nitrogen (linear correlation for 1988: $r=0.84$; 1989: $r=0.46$; 1990: $r=0.54$; $N=60$, $P<0.01$ for all). In 1988, the six monocultures on the richest soil mixture had dense litter layers (mean: 381 g m^{-2} of litter) and little living biomass (mean: 0.12 g m^{-2}). By 1990, they averaged 153 g m^{-2} of litter and 335 g m^{-2} of living biomass. In the two highest nitrogen soil mixtures, there were significant negative correlations between litter biomass one year and plant biomass the next ($r=-0.66$ and $r=-0.48$; $P<0.05$). Thus the dramatic decline in living biomass in the high-nitrogen plots from 1987 to 1988, and the suppression in 1989, may have been caused by accumulated litter.

Oscillations and chaos have been hypothesized to be rare for perennial plants because regulatory forces, such as competition for nutrients or light, act without a significant time delay. Litter causes a one-year time delay between growth and inhibition of future growth. Because litter production is greater in more productive plots, the magnitude of this time-delayed inhibitory force should increase with productivity. A simple model of plant-litter interactions exhibits bifurcations and chaos as soil nitrogen increases (Fig. 2*d*). Its two-point cycle, for instance, represents oscillation between a year of low litter and high living biomass, and a year of high litter and thus low living biomass.

Our results suggest that terrestrial plants may exhibit chaotic dynamics caused by a previously unsuspected mechanism, the inhibition of growth by accumulated litter. Other studies have reported litter inhibition of survival and growth, especially for herbaceous plants¹²⁻¹⁸. Such litter feedback could cause population oscillations and possibly chaos in productive habitats. Because litter is deposited where it is produced, however, such litter-driven chaotic dynamics may mainly occur at small spatial scales. Chaos may not be detectable on larger spatial scales if local dynamics are asynchronous or weakly coupled¹⁰. Moreover, unlike *Agrostis*, some plant species have robust shoots that can penetrate a deep litter layer¹². Such species should not

display chaotic dynamics by the mechanism described here and may displace species like *Agrostis* from productive habitats. Furthermore, both fire and grazing reduce litter mass^{17,18}, and should decrease the likelihood of chaotic dynamics in native grasslands. Native species that evolved with fire and grazing may be more likely to be inhibited by litter and displaced from habitats after fertilization, fire suppression, or cessation of grazing or mowing. Further work is needed to determine the generality of the patterns we report, to elucidate the role of chaotic dynamics versus stochastic factors in contributing to the fuzziness of patterns observed in nature, and to study the role of litter inhibition in structuring plant communities. □

Received 17 June; accepted 20 August 1991.

1. May, R. *Science* **196**, 645–647 (1974).
2. Li, T.-Y. & Yorke, J. A. *Am. Math. Monthly* **82**, 985–992 (1975).
3. May, R. & Oster, G. F. *Am. Nat.* **110**, 573–599 (1976).
4. Hassell, M. P., Lawton, J. H. & May, R. M. *J. Anim. Ecol.* **45**, 471–486 (1976).
5. Schaffer, W. M. *Am. Nat.* **124**, 798–820 (1984).
6. Schaffer, W. M. & Kot, M. In *Chaos: An Introduction* (ed. Holden, A. V.) 158–178 (Princeton University Press, Princeton, 1986).
7. Hassell, M. P. & May, R. M. in *Perspectives in Ecological Theory* (eds Roughgarden, J., May, R. M. & Levin, S. A.) 319–347 (Princeton University Press, Princeton, 1989).
8. Schaffer, W. M. & Kot, M. *J. theor. Biol.* **112**, 403–427 (1985).
9. Sugihara, G. & May, R. M. *Nature* **344**, 734–741 (1990).
10. Sugihara, G., Grenfell, B. & May, R. M. *Phil. Trans. R. Soc. B* **330**, 235–251 (1990).
11. May, R. M. *Proc. R. Soc. B* **228**, 241–266 (1986).
12. Sydes, C. & Grime, J. P. *J. Ecol.* **69**, 249–262 (1981).
13. Al Mufti, M. M., Sydes, C. L., Furness, S. B., Grime, J. P. & Band, S. R. *J. Ecol.* **65**, 759–791 (1977).
14. deJong, T. J. & Klinkhamer, P. G. I. *Oecologia* **65**, 153–160 (1985).
15. Bergelson, J. *Ecology* **71**, 2157–2165 (1990).
16. Carson, W. P. & Peterson, C. J. *Oecologia* **85**, 8–13 (1990).
17. Knapp, A. K. & Seastedt, T. R. *Bioscience* **36**, 662–668 (1986).
18. Bakker, J. P. *Nature Management by Grazing and Cutting* (Kluwer, London, 1989).
19. Wedin, D. & Tilman, D. *Oecologia* **84**, 433–441 (1990).

ACKNOWLEDGEMENTS. We thank A. El Haddi and C. Osbeck for assistance, our colleagues for comments, and NSF and the Mellon Foundation for financial support.

Retinal receptors in rodents maximally sensitive to ultraviolet light

Gerald H. Jacobs, Jay Neitz & Jess F. Deegan, II

Department of Psychology, University of California, Santa Barbara, California 93106, USA

HIGH sensitivity to near-ultraviolet light is a fundamental feature of vision in many invertebrates^{1,2}. Among vertebrates there are some amphibians, birds and fishes that are also sensitive to near-ultraviolet wavelengths^{3–6}. This sensitivity can be achieved through a class of cone photoreceptor containing an ultraviolet-sensitive pigment^{7–9}. Although these receptors were thought not to exist in the eyes of mammals, we now report that some rodents have a retinal mechanism that is maximally sensitive to ultraviolet light.

We measured the spectral sensitivity of pigmented house mice (*Mus musculus*) using a gross electrical potential, the electroretinogram (ERG). The electrical response to a flickering monochromatic test light sensed by an electrode placed against the surface of the eye was compared with that generated by an interleaved, similarly flickering, achromatic reference light¹⁰. Over successive stimulus presentations the intensity of the test light was varied until the response it produced just matched that produced by the fixed reference light. Repetition of this procedure for many test wavelengths can yield a reliable spectral sensitivity curve that accurately reflects the contributions from the various photoreceptor classes that operate under conditions of light adaptation¹¹.

Spectral sensitivity functions were obtained from five light-adapted mice (Fig. 1). There are two sensitivity maxima found at about 510 and 370 nm, suggesting that two distinct spectral

mechanisms are contributing to the potential under these conditions. If this is the case, then the substantial separation between the two peaks should enable the system responsible for the shorter of the two maxima to be isolated by chromatically adapting the eye. The ERG flicker photometric procedure was used to measure spectral sensitivity for five additional mice whose eyes were exposed to a bright orange adaptation light. Under these conditions sensitivity is sufficiently depressed at long wavelengths that ERG signals cannot be recorded to lights of wavelengths longer than about 440 nm. What remains after such adaptation are spectral sensitivity functions (Fig. 1, lower curve) that are virtually identical for all the mice—they peak at about 360 nm with a very rapid falloff (>0.3 log unit per 10 nm) in sensitivity towards the longer wavelengths.

The results of Fig. 1 indicate that spectral sensitivity in the light-adapted mouse probably results from the combined operation of two mechanisms, one giving peak sensitivity ($\lambda_{\max}$) at around 511 nm and the other in the ultraviolet. There are other explanations for this result⁵: for example, high ultraviolet sensitivity may represent the secondary absorbance peak (β band) of a photopigment having its principal peak in the visible spectrum; it may come from a photosensitive product which results from bleaching of the long-wavelength pigment, or it may arise because ultraviolet light causes some ocular structure(s) to fluoresce and this fluorescence could be emitted at a longer wavelength at which the pigment absorbing in the visible is sensitive. The relative heights of the two sensitivity peaks (Fig. 1, top) argue against the first explanation and the results of the chromatic adaptation experiment against the third. To test these three possibilities, we measured the sensitivity of the mouse ERG under each of four conditions. Thresholds were determined in two mice for 370- and 510-nm test lights flickering at 12.5 Hz

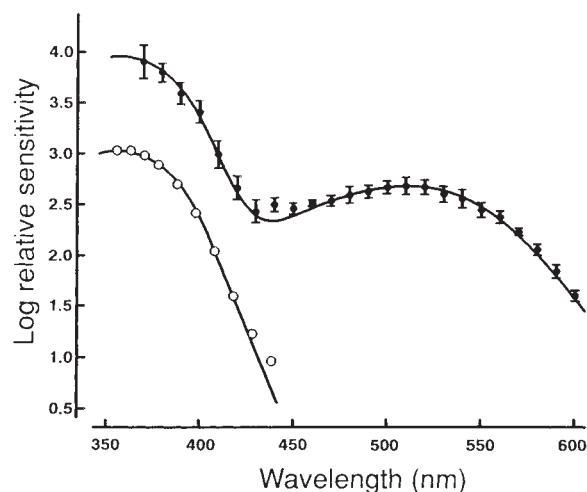


FIG. 1 Spectral sensitivity functions for mice obtained from ERG flicker photometric recordings. Top curve: solid circles are mean sensitivity values for 5 mice; error bars enclose ± 2 s.e. The test stimuli were monochromatic lights (half bandwidth 10 nm); the reference light was achromatic (corneal radiance, 0.05 mW). The flicker rate was 25 Hz. Ambient light in the room gave an additional constant luminance of 60 cd per m² at the cornea of the test eye. The solid line represents the best least-squares fit to the data of a template constructed by summing two photopigment curves ($\lambda_{\max}$ of 359 nm and 511 nm). These component curves were obtained from photopigment nomograms¹⁹ that had been translated along a log wavenumber axis²⁰ to the appropriate peak locations. Lower curve: mouse spectral sensitivity function measured while the eye was concurrently adapted to long-wavelength light (produced by a high-pass filter of 50% transmittance at 590 nm; corneal radiance, 6.2 mW). The adapting light (circular, 53°) was spatially coincident with the test and reference lights. Open circles represent mean values for 5 animals; the line is the best-fitting template curve derived as already described. The two sensitivity functions have been arbitrarily positioned on the ordinate.

in the presence of ultraviolet and yellow adaptation lights, the idea being that if there are two mechanisms that can be separately influenced, then sensitivity should alter more when the test and adaptation lights are from the same part of the spectrum than when they are not. That was the outcome: with 510 nm adaptation, the mean threshold change for the 510 nm test light was 1.19 log units, but the same adaptation had essentially no effect (average change of 0.01 log unit) on sensitivity to the 370 nm test light. Ultraviolet adaptation caused average threshold increases of 0.77 log unit (for 370 nm) and 0.37 log unit (for 510 nm). Note that in each case the homochromatic condition gives a greater loss in sensitivity than in the equivalent heterochromatic conditions and that there is no enhancement of sensitivity for heterochromatic adaptation, as would be expected if the bimodal spectral sensitivity function arose simply from a photoproduct. These results indicate that the spectral sensitivity function in the mouse depends on two photopic mechanisms—probably two types of cone—which are maximally sensitive in the near ultraviolet and at about 510 nm, respectively.

To see if the retinal mechanism sensitive to ultraviolet light might have some functional purpose, we tested mice in a behavioural test to see whether they could detect ultraviolet light¹². Spectral sensitivity functions for two mice are shown in Fig. 2. The functions are qualitatively similar to those from retinal recordings. The mice show two regions of peak sensitivity, one at about 510 nm and the other in the ultraviolet at the shortest wavelength at which it was possible to run the test (370 nm). These results establish that the ultraviolet-sensitive mechanism in the retina detected in electrophysiological measurements can contribute to vision in the mouse.

Is this ultraviolet-sensitive mechanism unique to the mouse? We recorded ERGs from three other rodent species which differ in their visual lifestyles and retinal construction—rat (*Rattus norvegicus*), gopher (*Thomomys bottae*) and gerbil (*Meriones unguiculatus*). These species have a different spectral position for the longer photopic system, but they all seem to show an ultraviolet response spectrally similar to the mouse (Fig. 3), although Fig. 3 indicates that its relative contribution to the ERG varies from species to species, perhaps because the relative proportions of sensitive cones can vary.

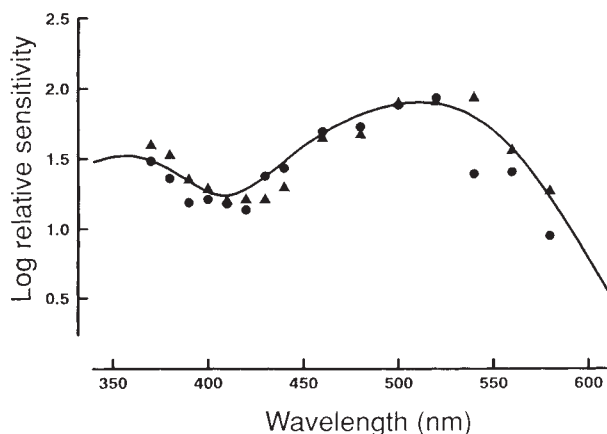


FIG. 2 Spectral sensitivity functions from mice in a behavioural test. Monochromatic stimuli were detected in a forced-choice discrimination task. The mice viewed three stimulus panels which were continuously illuminated with an achromatic light (4,800°; 8.2 cd per m²). On each trial a monochromatic test light was added to one panel. The animal's task was to indicate which of the panels had been illuminated by touching that panel. The test light was varied in wavelength and intensity over trials. The animal's ability to discriminate each of 15 test wavelengths was established by determining the intensity of the test light required to yield performance significantly better than chance ($P < 0.05$). The symbols are mean thresholds for each of two subjects. The line through the data is the best fit of the template function derived from summing two photopigment sensitivity curves (Fig. 1).

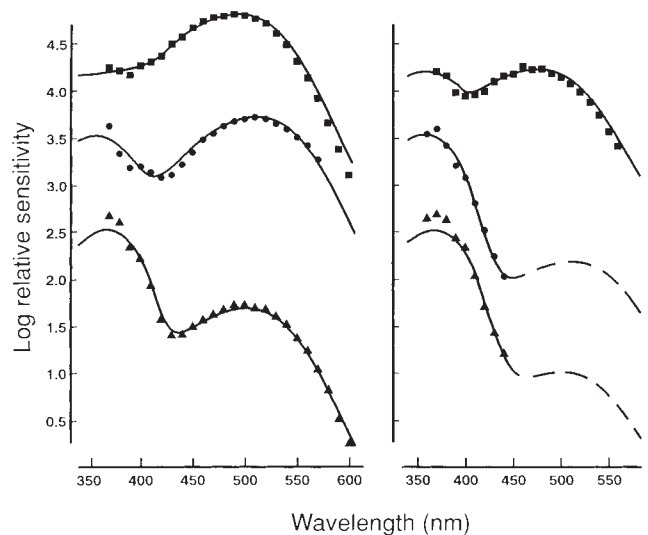


FIG. 3 ERG spectral sensitivity functions obtained from three rodent species (squares, gerbil; circles, rat; triangles, gopher) under test conditions involving achromatic adaptation (left) or long-wavelength adaptation (right). The adaptation conditions are described in Fig. 1. The templates best fit to the data were derived as for Fig. 1, except that for the gerbil the longer of the two nomogram photopigments had $\lambda_{\max}$ of 493 nm; for the gopher this pigment had $\lambda_{\max}$ of 500 nm. Functions are arbitrarily positioned on the ordinate.

Our experiments involve only a small sample of rodents, but an ultraviolet-sensitive mechanism may not be uncommon in the retinas of this group. This mechanism is maximally sensitive to ultraviolet light and enables short-wavelength light to be detected. Although the role(s) of this system in normal vision is unclear, it may explain certain other experimental observations^{13–15}. Earlier behavioural and electrophysiological experiments showed that rats¹⁶ and gerbils¹⁷ have only a single photopic mechanism active over the range from 440 to 600 nm, so they were accordingly classified as cone monochromats without colour vision. Our discovery of a second spectral mechanism maximally sensitive in the ultraviolet means that these rodents may have dichromatic colour vision. Other mammals have dichromatic colour vision which is based on signals from a short-wavelength cone ($\lambda_{\max}$ of 420–450 nm) and a cone type absorbing maximally at some longer wavelength¹⁸, so the rodent mechanism of colour vision may be unique among mammals. □

Received 28 May; accepted 27 August 1991.

1. Menzel, R. in *Handbook of Sensory Physiology* (ed. Autrum, H.), **7-6A**, 503–566 (Springer, Berlin, 1979).
2. Stark, W. S. & Tan, K. E. W. P. *Photochem. Photobiol.* **36**, 371–380 (1982).
3. Emmerton, J. & Delius, J. D. *J. comp. Physiol.* **141**, 47–52 (1980).
4. Goldsmith, T. H. *Science* **207**, 786–788 (1980).
5. Hawryshyn, C. W. & Beauchamp, R. D. *Vision Res.* **25**, 11–20 (1985).
6. Douglas, R. H. *J. comp. Physiol. A* **159**, 415–421 (1986).
7. Harosi, F. I. & Hashimoto, Y. *Science* **222**, 1021–1023 (1983).
8. Bowmaker, J. K. & Kunz, Y. W. *Vision Res.* **27**, 2102–2108 (1987).
9. Perry, R. J. & McNaughton, P. A. *J. Physiol.* **443**, 561–587 (1991).
10. Neitz, J. & Jacobs, G. H. *J. opt. Soc. Amer. A* **1**, 1175–1180 (1984).
11. Jacobs, G. H. & Neitz, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2545–2549 (1987).
12. Jacobs, G. H. *Vision Res.* **23**, 239–248 (1983).
13. Benshoff, H. M., Brainard, G. C., Rollag, M. D. & Lynch, G. R. *Brain Res.* **420**, 397–402 (1987).
14. van Norren, D. & Shellekens, P. *Vision Res.* **30**, 1517–1520 (1990).
15. Rapp, L. M., Toiman, B. L. & Dhindsa, H. S. *Invest. ophthalm. Vis. Sci.* **31**, 1187–1190 (1990).
16. Neitz, J. & Jacobs, G. H. *J. comp. Psychol.* **100**, 21–29 (1986).
17. Jacobs, G. H. & Neitz, J. *Experientia* **45**, 317–319 (1989).
18. Jacobs, G. H. *Comparative Color Vision* (Academic, New York, 1981).
19. Ebrey, T. G. & Honig, B. *Vision Res.* **27**, 147–151 (1977).
20. Baylor, D. A., Nunn, B. J. & Schnapf, J. L. *J. Physiol.* **390**, 145–160 (1987).

ACKNOWLEDGEMENTS. We thank J. Fenwick and K. Krogh for technical assistance. This work was supported by the NIH.

Cloned neuronal $I_K(A)$ channels reopen during recovery from inactivation

J. Peter Ruppersberg, R. Frank*, O. Pongs† & M. Stocker†

Max-Planck-Institut für Medizinische Forschung, Abteilung Zellphysiologie, Jahnstrasse 29, D-6900 Heidelberg, Germany

* Zentrum für Molekulare Biologie, D-6900 Heidelberg, Germany

† Lehrstuhl für Biochemie, Ruhr-Universität, D-4300 Bochum, Germany

THE kinetic behaviour and functional role of potassium ion (K^+) channels mediating a fast-inactivating K^+ current ($I_K(A)$) has been widely discussed¹. Activating in the subthreshold range of excitation, $I_K(A)$ channels are assumed to reduce the excitatory effect of depolarizing membrane currents in a time-dependent manner. Here we report that $I_K(A)$ channels not only open in response to a depolarization but open again after repolarization of the membrane. Although the current in response to the depolarization is rapidly inactivating, the current elicited by repolarization declines slowly and produces long-lasting afterhyperpolarizations under current-clamp conditions. This implies an additional physiological role for $I_K(A)$ channels, particularly those that activate positive to the threshold of excitation. The underlying biophysical mechanism was studied by fast-application of peptides corresponding to the N-terminal end of the $I_K(A)$ channel proteins. It was found to be a voltage-dependent release of the inactivation gate.

Three different mammalian $I_K(A)$ channels have been cloned, so far. RCK4 (ref. 2) and mShal (ref. 3) activate in the subthreshold range (around -60 mV) similar to the native $I_K(A)$ channels described¹. Raw3 (ref. 4) channels activate between -20 and 0 mV beyond the threshold of excitation. Raw3 channels therefore were not expected to influence excitability. Figure 1a shows that Raw3 channels do influence the membrane potential in the subthreshold range but only after a depolarization. This was measured in cell-attached patches under current-clamp conditions. Raw3 channels were activated by inward current pulses of -100 pA which decreased the membrane potential from any holding potential to about 0 mV, slightly positive of the activation threshold of Raw3 channels. Surprisingly, the membrane became hyperpolarized for several seconds after the depolarizing current pulses whenever the holding potential preceding the depolarization was positive to the K^+ equilibrium potential (around -90 mV). This suggests long-lasting channel activity of Raw3 channels at unusually negative potentials. Similar results were obtained in current-clamp experiments with RCK4 channels² and RCK4-1 channels⁵ (data not shown).

To visualize the afterhyperpolarization (AHP)-inducing channel activity at negative potentials we did voltage-clamp experiments on cell-attached patches with high K^+ (130 mM) in the patch pipette. Raw3 and RCK4 channels almost completely inactivate during a 1.5 -s depolarizing voltage pulse to 60 mV but reopen after repolarization to -80 mV giving rise to an inward current (Fig. 1b). This inward current is probably the cause of the AHP seen in current-clamp experiments (Fig. 1a).

The increase of channel activity after repolarization in the $I_K(A)$ channels seemed to be induced by the step in membrane voltage. It therefore might be caused by a voltage-dependent removal of inactivation. According to the ball and chain model of fast inactivation^{6,7}, the inactivation gate consists of the N-terminal end of the channel protein. This so-called 'ball domain' inactivates the channel in the open state by binding to its inner mouth, a process which up to now has been found to be independent of membrane voltage⁷. To test directly for voltage dependence of this process we applied peptides corresponding to the N-terminal end of $I_K(A)$ channels rapidly to inside-out patches with Raw3, RCK4 or RCK4-1 channels. In inside-out patches

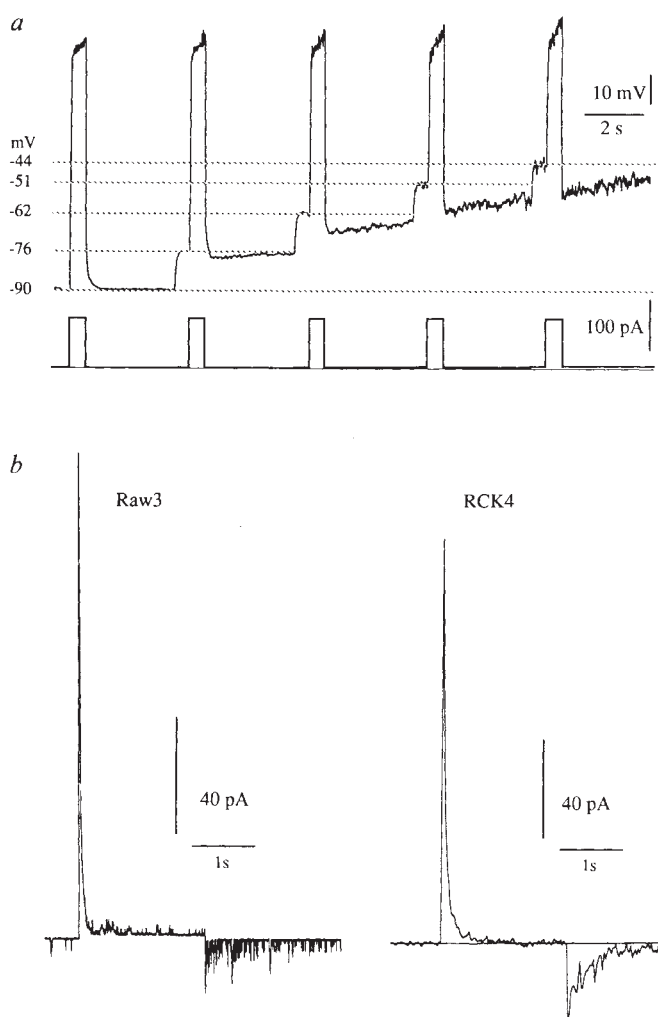
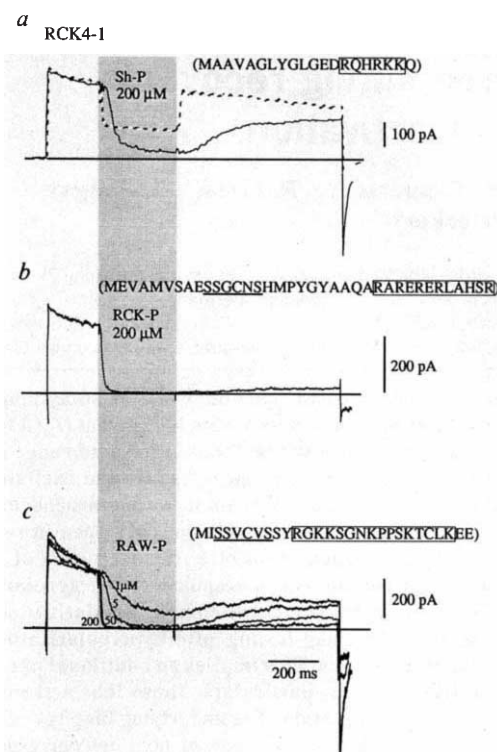


FIG. 1 Long-lasting AHPs are produced by openings of $I_K(A)$ channels at negative membrane potentials. **a**, Cell-attached patch with Raw3 channels measured under current-clamp conditions with normal frog Ringer's (NFR) solution in the patch pipette. With zero current the patch had a resting potential of around -50 mV (not shown). By applying an outward current of about 3 pA the potential was then brought to a prepotential of -90 mV. From this starting value a strong depolarization of 500 ms was provoked by an inward-directed clamp current of -100 pA. The resulting depolarization passed the activation threshold of Raw3 channels reaching about 0 mV. Preceding the following depolarizations the holding current was decreased in steps of 1 pA leading to prepotentials of -76 , -62 , -51 and -44 mV. At the end of the depolarization the potential either reached its starting value again (at a prepotential of -90 mV) or hyperpolarized transiently (at -76 – -44 mV). **b**, Cell-attached patches with Raw3 and RCK4 channels measured under voltage-clamp conditions with K^+ frog Ringer's (KFR) solution in the pipette. Test pulses going from -80 mV to 60 mV eliciting a transient K^+ outward current. At the end of the test pulse of 1.5 s the channel activity almost decreased to zero. Channels reopen after the repolarizing voltage step to -80 mV. Channel activity after repolarization is markedly higher than at the end of the test pulse.

METHODS. Capped run-off complementary RNA of Raw3, RCK4 and RCK4-1 was synthesized using a standard protocol and injected into *Xenopus* oocytes. Patch-clamp recordings were made 2–7 days after injection. NFR solution contained (in mM): NaCl, 115; KCl, 2.5; CaCl₂, 2; HEPES, 10 (pH 7.2, NaOH). KFR solution was as NFR but sodium was replaced with potassium. Pipettes made from thick-wall borosilicate glass were filled with either NFR or KFR and had a resistance of 4 M Ω and a tip diameter of about 2 μ m. Currents were recorded with an EPC9 amplifier (Heka Electronic, Lamprecht, Germany). In voltage-clamp experiments leakage and capacitive currents were measured and compensated linearly with the EPC9 at -80 mV membrane potential. Evaluation of data and fitting of exponentials was done off-line with an IBM AT Computer.

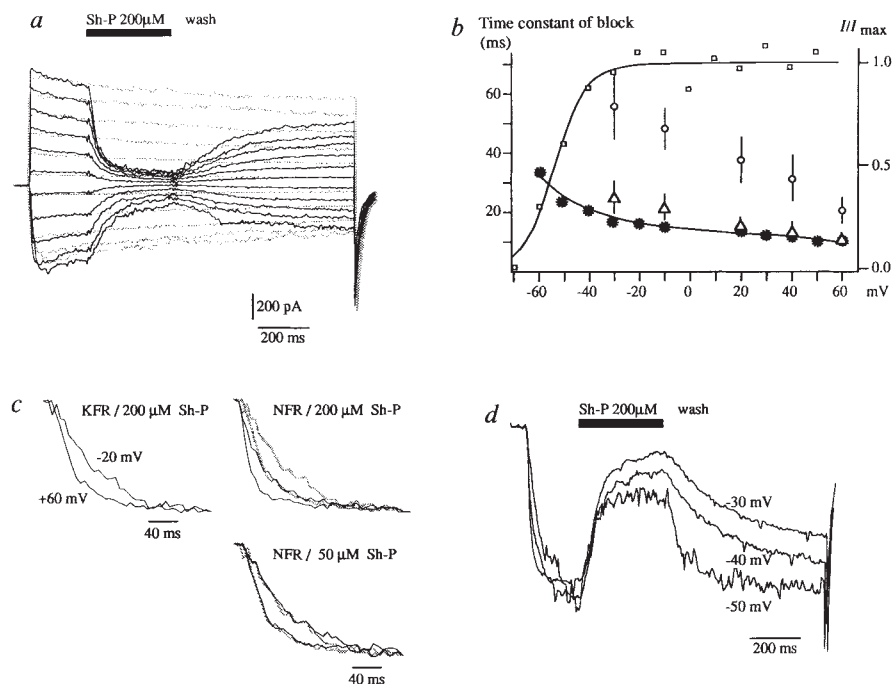
FIG. 2 Fast-application of three different ball peptides to inside-out patches containing RCK4-1 channels in symmetrical K^+ solution. K^+ outward currents were recorded in response to steps in membrane potential (test pulses) of 1.2 s from -80 to $+40$ mV. After 250 ms of onset of the test pulse the patch was moved for 300 ms from K-Int solution into K-Int solution containing the respective ball peptide. The amino-acid sequence of the peptides is indicated in the insets. The subdomain of the charged residues is marked with a box, the putative regulation site⁵ is underlined. *a*, Binding and unbinding of Sh-P (first 22 residues of fast-inactivating *Drosophila Shaker* K^+ channels). Binding is faster (31 ± 8 ms; $n=6$) than unbinding (200–300 ms). Accordingly there is a steady-state current of about 13% flowing in presence of $200 \mu M$ Sh-P. The dotted trace represents the speed of the fast-application system. It was measured by fast application of K-Int solution containing 30 mM Cs producing a fast channel block of about 50% of the channels. There is a delay between the Cs block and the onset of the peptide action which might reflect an access step for the peptides on the inside of the channel. *b*, The N-terminal peptide of RCK4 (RCK-P) consisting of the first 37 residues of the RCK4 sequence binds with a time constant of 13 ± 4 ms ($n=4$) and takes seconds to washout. *c*, The N-terminal peptide of Raw3 (Raw-P; first 28 residues of the Raw3 sequence) in $200 \mu M$ concentration binds so fast (3.7 ms) that the binding rate could not be separated from the speed of the application system. In lower concentrations the binding rate was: 6 ms for $50 \mu M$, 24 ms for $5 \mu M$ and 54 ms for $1 \mu M$. Washout of the peptide also depends on the concentration. A possible reason for this might be unspecific binding of the peptide. Traces are normalized to the current value immediately before application of the peptide. The scale bar refers to the experiment with $50 \mu M$ Raw-P in which the patch was first exposed to Raw-P. In the following experiments with 5, 1 and $200 \mu M$ Raw-P the channels showed some fast inactivation before fast-application of Raw-P. This artifact was due to the incomplete washout of Raw-P.

METHODS. Peptides were synthesized in a continuous flow instrument constructed and operated as described in ref. 12. Peptide chain assembly was performed on a 1% crosslinked polystyrene support using Fmoc chemistry and *in situ* activation by BOP (ref. 13). Purification was done by reversed-phase HPLC. Excised inside-out patches were kept in K-Int solution



containing (in mM): KCl, 100; $MgCl_2$, 2; EGTA, 10; HEPES, 10 (pH 7.2, KOH, 20). K-Int solution was put into the recording chamber after some minutes of vortexing in air to allow equilibration of the oxygen partial pressure. Fast application of the ball peptides was done as described¹⁴.

FIG. 3 Voltage-dependence of block of RCK4-1 channels by Sh-P. *a*, Family of current responses to depolarizing test potentials between -40 and 60 mV in 10 mV increments measured with KFR in the recording pipette. Without Sh-P (dotted traces) currents mediated by RCK4-1 channels show slow inactivation to about 86% of their maximum during the test pulse of 1.2 s. Sh-P blocks both inward and outward currents. *b*, Voltage dependence of the binding rate of $200 \mu M$ Sh-P in symmetrical K^+ solution ($\circ$) obtained from six patches and with NFR in the pipette (Δ) obtained from five patches. Standard deviations of the fitted time constants are given by the bars. The voltage dependence of the time constant of block was about 90 mV for an e-fold change for both NFR and KFR in the peptide. Considering four charged amino-acid residues in the peptide, this corresponds to 7% of the membrane potential being passed by the ball peptide to reach the receptor. The voltage dependence of the binding rate was determined in the range where the steady-state activation curve saturated. The steady-state activation curve ($\square$) was obtained by measuring the tail current amplitudes from the control experiment (dotted traces) in *a*. Tail currents elicited by the back-step to -80 mV were normalized to their maximum value and fitted with a Boltzmann isotherm yielding a half maximal activation at -54 mV. The steady-state block by $200 \mu M$ Sh-P measured in symmetrical K^+ (stars, line drawn by hand) was also voltage-dependent. *c*, Direct comparison of normalized time courses of binding Sh-P at two different membrane potentials (-20 and 60 mV). Traces are maximum-scaled to the current value before applying the peptide and minimum-scaled to the current after 240 ms of peptide action. The left traces are obtained from the measurement in *a* and are again plotted for comparison (dotted traces) on the right side (time constants were 33 and 75 ms). On the right side the upper traces show the increase in binding rate after reduction of



K^+ (130 mM to 2.5 mM) on the extracellular side of the patch (13.5 and 33 ms at 60 and -20 mV, respectively). The lower traces show again an experiment with 2.5 mM K^+ but with a reduced concentration of $50 \mu M$ Sh-P (time constants were 34 and 82 ms). *d*, Current responses at negative potentials scaled to their maximum. The rate of unbinding of Sh-P becomes faster at very negative potentials. At -50 mV it is as fast as the rate of binding leading to a steady-state current in the presence of the peptide of almost 50% of the peak current.

these $I_K(A)$ channels lose their fast inactivation by an oxidative process⁵. N-terminal peptides should then act as pore-blocking 'ball-peptides' and restore fast inactivation as shown for the N-terminal peptide of *Drosophila Shaker* channels⁶. We used the N-terminal 'ball peptides' from Raw3 (Raw-P), *Shaker* (Sh-P) and RCK4 (RCK-P, which is also the N-terminal peptide of RCK4-1) channels (sequences given in the insets in Fig. 2). These peptides were each blocking all three channel types (Raw3, RCK4 and RCK4-1). But neither Raw3 nor RCK4 channels were well suited for fast-application experiments. Raw3 channels had a quick run-down in inside-out patches and RCK4 channels had a slow inactivation in the time range of hundreds of milliseconds. We therefore used RCK4-1 channels which have no run-down and a slow inactivation in the range of tens of seconds⁷.

Figure 2 shows the effect of 200 μ M Sh-P, 200 μ M RCK-P and 1, 5, 50 and 200 μ M Raw-P on RCK4-1 channels at +40 mV. Raw-P had the strongest blocking effect whereas Sh-P was unbinding most rapidly. Because of its fast washout we used Sh-P for experiments with a series of pulses of different voltages which require a complete washout between the individual pulses. Figure 3a shows families of inward and outward currents mediated by RCK4-1 channels with (solid trace) and without (dashed traces) rapid application of 200 μ M Sh-P measured in symmetrical K^+ solution (130 mM). The time constant of binding (Fig. 3b, circles; Fig. 3c, left traces) as well as the steady-state current (Fig. 3b, triangles) in presence of Sh-P depended on membrane voltage. The voltage dependence of the binding rate was not related to the steady-state activation curve of the RCK4-1 channels (Fig. 3b, squares) which saturated at -30 mV. The binding rate was markedly faster but still voltage-dependent when the extracellular K^+ was lowered to 2.5 mM (Fig. 3b, triangles; Fig. 3c, right traces). In 2.5 mM K^+ , reduction of the concentration of Sh-P to 50 μ M (Fig. 3c, lower traces) results in the same binding rates as in 130 mM K^+ with 200 μ M Sh-P. This indicates that the influx of K^+ interacts with the binding of Sh-P but is not responsible for its voltage dependence.

For negative voltages down to -40 mV the time constant of unbinding becomes as fast as the time constant of binding (Fig. 3d; traces are scaled to their maximum value). Similar experiments were done with Raw3 channels and Raw-P and with RCK4 channels and RCK-P (not shown). Although these results were less clear they never contradicted those shown in Fig. 3.

The voltage dependence of binding and unbinding of the N-terminal peptide is probably the reason why reopenings occur after a repolarizing voltage step. But it does not explain why these reopenings persist over such a long time, so why is the reopening of the channels in the negative voltage range long-lasting and not terminated by the faster process of deactivation (transition back from the open to the activatable closed state)? The experiments shown in Fig. 4 address this question. Here we used a high concentration (200 μ M) of Raw-P because of its high potency of blocking the RCK4-1 channels. In Fig. 4a Raw-P was only applied briefly. In response to the repolarization a large tail current of deactivation (time constant of 118 ms) is visible because of the high extracellular K^+ concentration. When Raw-P is applied but not washed out again (Fig. 4b, repolarization also produces a tail current although the outward-current was fully blocked. This tail current probably reflects voltage dependence of block by the peptide as shown for Sh-P in Fig. 3. The much slower decline of the tail current in presence of Raw-P (730 ms) suggests that the time course of this tail current is controlled by reopening of channels rather than by deactivation. This also suggests that the channels are not able to deactivate while the ball peptide is bound. In Fig. 4c Raw-P was washed out during the decaying phase of the tail current. The wash out produced an additional increase of the inward current superimposed on the tail current. This demonstrates that the inward current is really mediated by channels which unbind the peptide.

In contrast to the decay time constant, the area under the tail current was almost the same in Fig. 4a, b and c and even slightly larger in the presence of the peptide ($112 \pm 4.4\%$, mean and standard deviation, $n = 3$). This is only possible if the channels are unable to deactivate while being blocked by the ball peptide. We conclude that channels must unbind the peptide before they deactivate.

Fast inactivation by the N-terminal ball peptides depends on membrane voltage. It prevents deactivation of the K^+ channel as long as the ball peptide is bound to the channel. Unbinding of the ball peptide is also voltage dependent and leads to reopening of the channel at negative potentials. As recovery from fast inactivation (unbinding of the peptide) is a slow process in RCK4 and Raw3 channels, those reopenings take several seconds. Thus, $I_K(A)$ channels may produce long

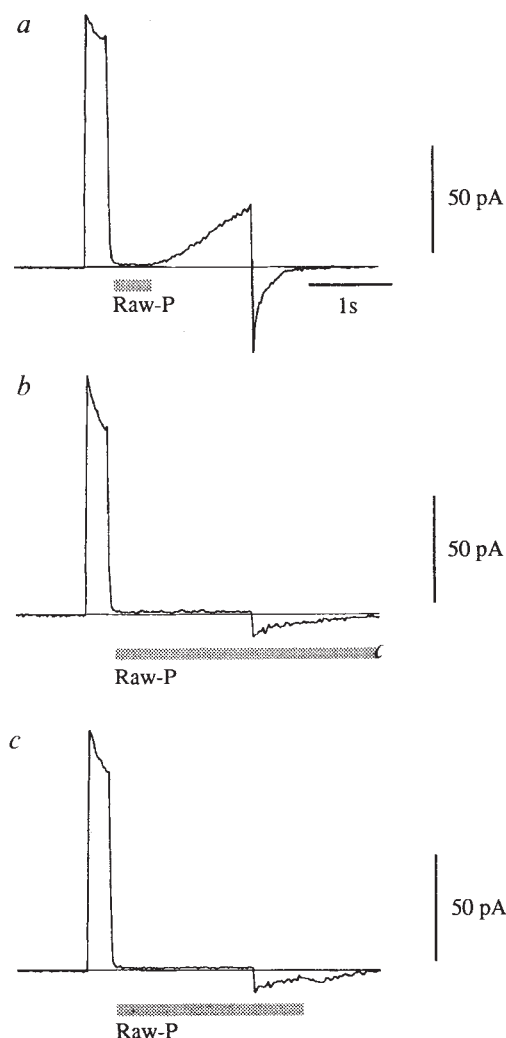


FIG. 4 Reopening of K^+ channels inactivated by a ball peptide. Test pulses of 2 s from -80 mV to 50 mV and back to -80 mV. RCK4-1 channels as in Figs 2 and 3. The current was measured in symmetrical K^+ (pipette filled with KFR) to visualize channel openings at -80 mV as inward tail currents. a, Raw-P is applied for 500 ms during the test pulse to show the time courses of binding and unbinding. The tail current after the step back to -80 mV declines rapidly because of channel deactivation (transition from the open to the closed state) and has a time constant of 118 ms. b, Test pulse as in a but continuous application of Raw-P. Although almost all channels are blocked by 200 μ M Raw-P at 50 mV, a considerable inward tail current is elicited by the step back to -80 mV. The decline of the tail current in presence of Raw-P (730 ms) is slower than deactivation. c, Experiment as in b but washout of Raw-P during the tail current leading to an additional increase of the inward current.

hyperpolarizing afterpotentials although they have been completely inactivated by a preceding depolarization. This might be an important second task of $I_K(A)$ channels and would reduce the frequency of action potentials by producing a long-lasting AHP.

Long-lasting AHPs have been observed in several types of neurons⁸⁻¹¹. In almost all cases the most probable explanation was an activation of calcium-dependent K^+ channels. From our data it is evident, however, that A-type K^+ channels can also be responsible for a long-lasting increase of K^+ conductance after depolarization. □

Received 10 June; accepted 15 August 1991.

1. Rudy, B. *Neuroscience* **25**, 729-750 (1988).
2. Stühmer, W. et al. *EMBO J.* **8**, 3235-3244 (1989).
3. Pak, M. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4386-4390 (1991).
4. Schröter, K. H. et al. *FEBS Lett.* **278**, 211-216 (1991).
5. Ruppersberg, J. P. et al. *Nature* **352**, 711-714 (1991).
6. Zagotta, W. N., Hoshi, T. & Aldrich, R. W. *Science* **250**, 568-570 (1990).
7. Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Science* **250**, 533-538 (1990).
8. Lancaster, B. & Wheal, H. V. *Neuroscience* **127**, 267-275 (1984).
9. McCormick, D. A., Connors, B. W., Lighthall, J. W. & Prince, D. A. *J. Neurophysiol.* **54**, 782-805 (1985).
10. Lancaster, B. & Adams, P. R. *J. Neurophysiol.* **55**, 1268-1282 (1986).
11. Lancaster, B. & Nicoll, R. *J. Physiol.* **389**, 187-203 (1986).
12. Frank, R. & Gausepohl, H. in *Modern Methods in Protein Chemistry* Vol. 3 (ed. Tschesche, H.) 42-60 (de Gruyter, Berlin and New York, 1988).
13. Castro, B., Domroy, J. R., Evin, G. & Selve, C. *J. chem. Res.* **182**, 2118-2122 (1977).
14. Franke, C. H., Hatt, H. & Dudel, J. *Neurosci. Lett.* **77**, 199-204 (1987).

ACKNOWLEDGEMENTS. We thank B. Sakmann, T. Verdoorn, P. Jonas, P. Stern, J. Cox, G. Stuart, M. Koenen and K. H. Schröter for helpful discussions and for reading the manuscript and B. Werner and A. Bürk for secretarial assistance. This work was supported by the Deutsche Forschungsgemeinschaft.

On the complexity of self

Alexander Yu. Rudensky, Satyajit Rath*,
Paula Preston-Hurlburt, Donal B. Murphy†
& Charles A. Janeway Jr.‡

The Section of Immunobiology, Yale University School of Medicine and
The Howard Hughes Medical Institute, New Haven,
Connecticut 06510, USA

SELF peptides bound to self major histocompatibility complex (MHC) molecules have been implicated both in positive^{1,2} and in negative^{3,4} selection of T cells during intrathymic development. We report here that the novel MHC-restricted monoclonal antibody Y-Ae (ref. 5) detects the MHC class II bound form of a major self peptide⁶. Y-Ae binds ~12% of the relevant MHC class II molecules on self antigen presenting cells. The peptide detected by Y-Ae is one of several major peptides eluted from the MHC molecule⁶. These data suggest that self peptides presented by self MHC class II molecules at densities sufficient to signal a CD4 T cell are of very limited complexity. Furthermore, as Y-Ae stains antigen presenting cells that mediate negative selection but not thymic cortical epithelial cells⁵ that drive positive selection³, differential expression of self peptide:self MHC class II complexes may be a key feature of intrathymic selection.

The Y-Ae monoclonal antibody precipitates a fraction of I-A^b molecules from spleen cells of mice that express both surface I-A^b and I-E^b molecules. By immunofluorescence or quantitative binding analysis⁵, about 12% of surface I-A^b molecules on B cells from strain B10.A(5R) (I-A^b, I-E⁺) stain with Y-Ae (Fig. 1a), whereas no staining is observed on B cells from strain C57BL/6J(B6) (I-A^b, I-E⁻) (Fig. 1b). When I-A^b molecules were isolated from the Y-Ae-positive B lymphoma cell line LB27.4 (Fig. 1c), one of six major RP-HPLC peaks⁶ analysed contained a peptide representing residues 56-73 of the I-E α

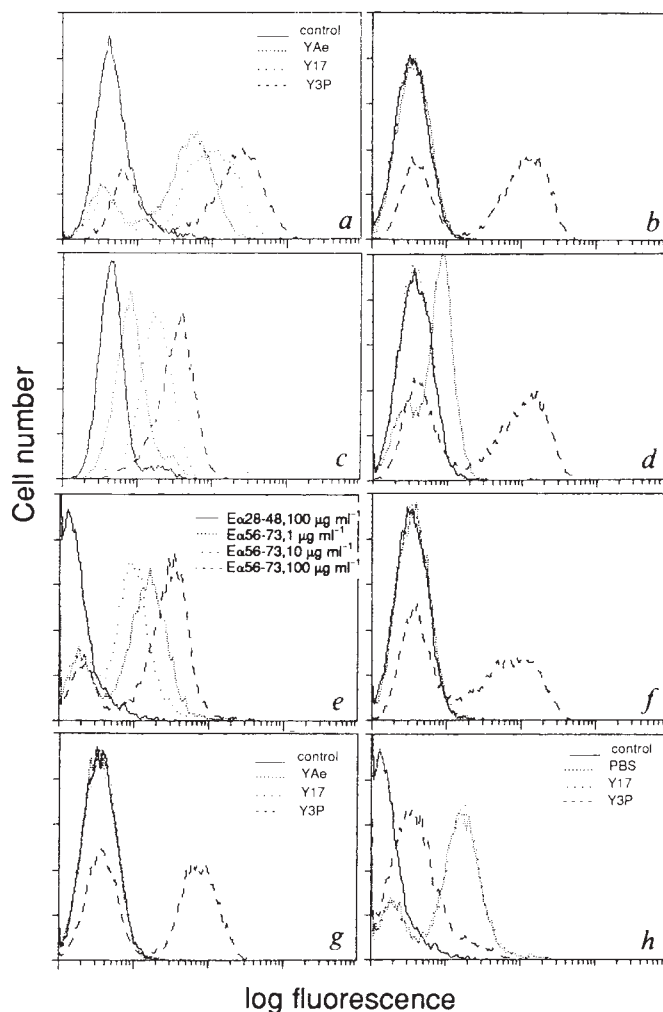


FIG. 1. Eα 56-73 peptide bound to I-A^b is recognized by the Y-Ae monoclonal antibody. Negative control profiles are shown as solid lines. Surface staining of B10.A(5R) (a) and C57BL/6J (B6) (b) spleen cells, B-lymphoma cell line LB27.4 (ref. 9) (c) with the anti-I-A^b antibody Y-3P (ref. 10), anti-I-E^b antibody Y-17 (ref. 11) and Y-Ae antibody⁵. Staining of B6 spleen cells pulsed with 20 μg ml⁻¹ Eα 56-73 peptide (d) and Eα 28-48 peptide (f) with the anti-I-A^b antibody Y-3P, anti-I-E^b antibody Y-17 and Y-Ae antibody. e, Y-Ae antibody binding to B6 spleen cells pulsed with Eα 28-48 peptide, 100 μg ml⁻¹, Eα 56-73 peptide, 1, 10 and 100 μg ml⁻¹. g, Staining of B6-H-2^{bm12} spleen cells pulsed with 20 μg ml⁻¹ Eα 56-73 peptide with the anti-I-A^b antibody Y-3P, anti-I-E^b antibody Y-17 and Y-Ae antibody. h, Binding of Eα 56-73 peptide, 10 μg ml⁻¹ to B6 spleen cells in the absence of inhibitors, or in the presence of 5 μg ml⁻¹ Y-17 or Y-3P antibodies.

METHODS. Viable spleen cells were isolated from suspensions by passage over a Ficoll-Hypaque density cushion ($d=1.077$, lymphocyte separation medium, Organon) and washed in PBS, pH 7.2, and stained. The peptides Eα 56-73 (ASFEAQGALANIADVKA; single-letter amino-acid code) and Eα 28-48 (FDGDEIFHVDIEKSETIWR) were synthesized using an Applied Biosystems Automated Peptide Synthesizer, purified by HPLC, and confirmed by mass spectroscopy. In peptide-pulsing experiments spleen cells were incubated with various doses of peptides in Clock's medium (Irvine) containing 5% FCS, 3×10^{-5} M 2-mercaptoethanol, 2×10^{-4} M L-glutamine for 5 h at 37 °C in the absence or presence of 5 μg ml⁻¹ blocking antibody. All staining was done in 96-well plates on $2-10 \times 10^5$ cells. Primary antibodies ($1-2 \mu\text{g ml}^{-1}$) were appropriately diluted in PBS containing 1% FCS and 0.1% sodium azide (staining buffer) and added to the cells (100 μl). The cells were incubated with the primary antibodies at 4 °C for 30-60 min, washed with staining buffer and similarly treated with rabbit anti-mouse IgG-Fc-fluorescein (Accurate). In antibody blocking of Eα 56-73 peptide binding to B6 spleen cells experiments, Y-Ae was used as biotinylated antibody and detected by avidin D-fluorescein (Vector). Fluorescent cytometry was carried out on a FACScan flow cytometer (Becton-Dickinson).

* Present address: National Institute of Immunology, JNU Complex, Shahid Jeet Singh Road, New Delhi 110 067, India.

† Present address: Laboratory for Immunology, D421, Wadsworth Center, New York State Department of Health, Empire State Plaza, Albany, New York 12201, USA.

‡ To whom correspondence should be addressed.

chain (E α 56-73). This suggested that E α 56-73 bound to I-A^b might be the determinant recognized by Y-Ae. To test this, E α 56-73 was synthesized and added to B cells from B6 mice; these cells then stained strongly with Y-Ae (Fig. 1d). Binding is time, temperature and dose dependent (Fig. 1e; data not shown). Similar staining is not obtained with other peptides (data not shown), including the E α 28-48 peptide (Fig. 1f). Y-Ae stains B cells of I-E⁺ strains that express I-A^b, but not B cells of I-E⁺ mice that express other I-A alleles, including the mutant strain B6-H-2^{bm12}, which differs from B6 only by three amino-acid residues in the A β chain. Similarly, addition of E α 56-73 to B cells of strain B6-H-2^{bm12}, or strains of other MHC haplotypes (data not shown), does not elicit staining by Y-Ae (Fig. 1g). Finally, monoclonal antibodies specific for I-A^b, but not for I-E, inhibit Y-Ae binding to B6 spleen cells incubated with E α 56-73, whether they are added during (Fig. 1h) or after incubation with E α 56-73 (not shown). Thus, it is clear that Y-Ae recognizes a complex of E α 56-73 with I-A^b.

In addition to surface staining by Y-Ae, the dominance of E α 56-73 bound to I-A^b is seen in two other ways. First, Y-Ae reacts with a fraction of I-A^b molecules isolated from B10.A(5R) LPS-activated B cells, as I-A^b molecules can be separated on Y-Ae sepharose into Y-Ae⁺ (10-15%) and Y-Ae⁻ (85-90%) fractions, both of which label equally with anti-I-A^b antibodies on western blots (Fig. 2). Second, integration of the RP-HPLC profile of peptides eluted from I-A^b molecules isolated from LB27.4 indicates that E α 56-73 represents 5.7% of the peptides isolated⁶, whereas Y-Ae reacts with 8% of cell-surface I-A^b on LB27.4 (Fig. 1c). This difference in yield may reflect the fact

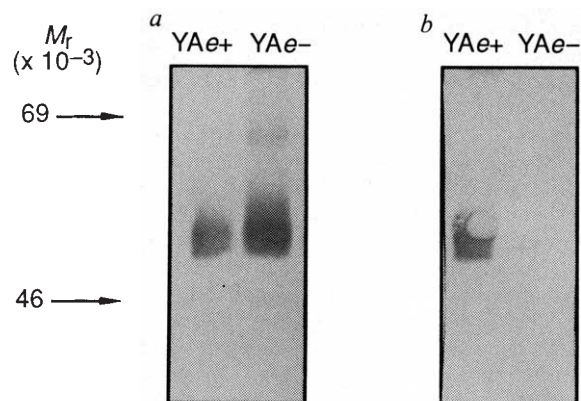
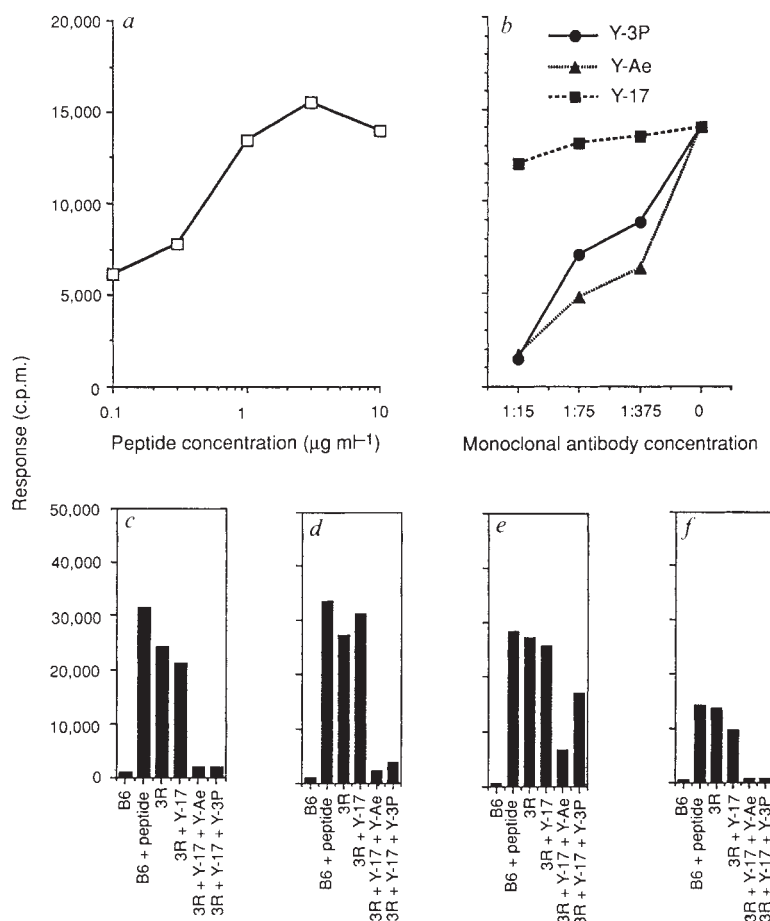


FIG. 2 Western blot analysis of purified (Y-Ae⁺) and (Y-Ae⁻) I-A^b molecules using anti-I-A^b Y-3P antibody (a) and Y-Ae antibody (b). Only Y-Ae⁺ material reacts with Y-Ae, indicating purity, whereas Y-3P reacts with both preparations. Y-Ae and Y-3P react equally with Y-Ae⁺ material showing that both antibodies bind quantitatively to this protein complex. Relative molecular mass of protein standards is indicated as $M_r \times 10^{-3}$ on the left.

METHODS. I-A^b molecules were purified from cell lysate of LPS blasts from B10.A(5R) spleen cells by affinity chromatography on Y-3P column as described⁶. Purified I-A^b molecules were similarly further fractionated on a Y-Ae column into Y-Ae⁺ and Y-Ae⁻ fractions and analysed by 10% SDS-PAGE gel run under nonreducing conditions and immunoblotted with Y-3P and Y-Ae antibodies. Horseradish peroxidase-conjugated rabbit anti-mouse IgG was used as a secondary antibody.

FIG. 3 T cells from C57BL/6 (B6) mice respond to E α 56-73 presented by I-A^b and crossreact with a Y-Ae-inhibitable epitope on B10.A(3R) spleen cells. a, Lymph node cells from B6 mice primed with E α 56-73 proliferate in response to E α 56-73 (Δ c.p.m., background of 7,000 c.p.m. subtracted from all groups). Similarly immunized 5R mice did not respond above background (data not shown). b, Inhibition of the lymph node proliferative response to $10 \mu\text{g ml}^{-1}$ E α 56-73 by monoclonal antibodies to I-A^b (Y-3P) or I-E^b (Y-17), or by Y-Ae, assayed as in a. c-f, Interleukin-2 production by individual T-cell hybrids derived from B6 T cells in response to E α 56-73 presented by B6 spleen cells or to 3R spleen cells incubated alone or with anti-I-E^b (Y-17) alone or with added anti-I-A^b (Y-3P) or Y-Ae.

METHODS. B6 and 5R mice (8-10-week-old) were immunized with $20 \mu\text{g}$ E α 56-73 peptide in complete Freund's adjuvant into the hind footpads. On day 8 after immunization lymph node cells were analysed for proliferative responses to the peptide by incubating 4×10^5 cells per well in flat-bottom 96-well plates in the presence of serial dilutions of peptide in Click's medium (Irvine) containing 5% FCS, 3×10^{-5} M 2-mercaptoethanol, 2×10^{-4} M L-glutamine and antibiotics for 80 h at 37°C , 5% CO_2 . Cells were pulsed with $1 \mu\text{Ci}$ per well of ^3H -thymidine for the last 18 h of culture, collected and counted for thymidine incorporation (Betaplate, LKB). Background c.p.m. = 7,000 in the absence of antigen have been subtracted from all groups to give Δ c.p.m. In the antibody blocking experiments, lymph node cells were preincubated with the dilutions of culture supernatants from Y-3P, Y-17 and Y-Ae hybridomas for 30 min before peptide was added to the wells. E α 56-73 peptide-specific hybrid T-cell lines were generated by fusion of E α 56-73-activated T-cell blasts from B6 mice with a T-cell receptor chain-negative BW5147 thymoma. Stimulation of the hybrid T-cell lines was measured by incubating 1×10^5 hybrid cells with 1×10^5 antigen-presenting cells per well of 96-well plates in the presence or absence of peptide. After 24 h of incubation $50 \mu\text{l}$ aliquots of supernatants were removed and tested for lymphokine using the CTLL-2 cell line at 5,000 cells per well in $100 \mu\text{l}$ cultures for 24 h. Cells were pulsed for 4 h with $1 \mu\text{Ci}$ per well of ^3H -thymidine and collected. In the antibody blocking experiments with T-cell hybrids, APC were preincubated for 30 min with the culture supernatants



from Y-17 alone or together with Y-3P or Y-Ae (1:8 dilution) before T cells were added to the wells.

that cytoplasmic forms representing about 30% of total I-A^b molecules are Y-Ae negative (S.R., A.Y.R. and C.A.J., unpublished data). Thus, E α 56-73 is a major self peptide associated with I-A^b.

If E α 56-73 is an important self peptide presented by I-A^b in I-E⁺ strains of mice, then T cells from these mice should be tolerant to this peptide. To examine this, B6 and 5R mice were immunized with E α 56-73. Although 5R mice do not respond (not shown), B6 mice make a strong proliferative response to E α 56-73 (Fig. 3a). The B6 response to E α 56-73 is inhibitable by Y-Ae and anti-I-A^b, but not by anti-I-E^b (Fig. 3b). Several cloned T-cell hybrids prepared from B6 T cells responding to E α 56-73 also respond strongly to spleen cells from the MHC-identical strains B10.A(3R) (Fig. 3c-f) and 5R (data not shown). Their response to 3R spleen cells is inhibited by anti-I-A^b and by Y-Ae, but not by anti-I-E^b monoclonal antibody (Fig. 3c-f). Thus, 3R and 5R mice express E α 56-73 bound to I-A^b and are tolerant to it. These data further confirm binding of E α 56-73 to the peptide-binding cleft of I-A^b molecules.

Our current⁶ and previous⁵ data on Y-Ae suggest that immunological self may be of limited complexity. Y-Ae reacts with the 12% of I-A^b molecules binding E α 56-73 on antigen-presenting cells in strain B10.A(5R)⁵. This peptide is one of four to five defined major peptides eluted from I-A^b molecules that

together seem to represent 50% of the peptides⁶ isolated from a B-lymphoma line in association with I-A^b. As Demotz *et al.*⁷ and Harding and Unanue⁸ have shown that 200 (or ~1%) of the MHC molecules on normal antigen-presenting cells need to be occupied by a specific peptide to activate CD4 T-cell hybrids, a maximum of 100 different peptides can be presented at any one time. Furthermore, even higher levels of self peptide: self MHC complexes may be required to activate naive T cells⁷, which generally require more T cell receptor ligation for activation. If many of the available MHC class II molecules are occupied by a few endogenously synthesized peptides, as our data suggest, then the total number of distinct self peptide: self MHC complexes present at levels sufficient to activate a naive CD4 T cell may be quite small. Further studies of the type reported here and in ref. 6 will be required to fully define the complexity of self peptides presented by self MHC molecules. Nevertheless, our data indicate that self may be even less complex than could be inferred from constraints imposed by the data of Demotz *et al.*⁷ and of Harding and Unanue⁸.

These data have interesting implications for positive selection, which occurs on Y-Ae negative thymic cortical epithelium, and for negative selection, which occurs on Y-Ae-positive antigen-presenting cells presenting peptides of limited complexity. A full discussion of these implications will be published elsewhere (A.Y.R. *et al.*, manuscript in preparation). □

Received 6 August; accepted 18 September 1991.

1. Nikolic-Zugic, J. & Bevan, M. J. *Nature* **344**, 65-67 (1990).
2. Berg, L., Frank, G. D. & Davis, M. M. *Cell* **60**, 1043-1052 (1990).
3. vonBoehmer, H., Teh, H. S. & Kisielow, P. *Immun. Today* **10**, 61-66 (1989).
4. Murphy, K. M., Meimberger, A. B. & Loh, D. Y. *Science* **250**, 1720-1723 (1990).
5. Murphy, D. B. *et al.* *Nature* **338**, 765-767 (1989).
6. Rudensky, A. Y., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. & Janeway, C. A. Jr *Nature* **353**, 622-628.
7. Demotz, S., Grey, H. M. & Sette, A. *Science* **249**, 1028-1030 (1990).

8. Harding, C. V. & Unanue, E. R. *Nature* **346**, 574-576 (1990).
9. Kappler, J., White, J., Wegmann, D., Mustain, E. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3604-3607 (1982).
10. Janeway, C. A. Jr *et al.* *J. Immun.* **132**, 662-669 (1984).
11. Lerner, E. A. *et al.* *J. exp. Med.* **152**, 1085-1101 (1980).

ACKNOWLEDGEMENTS. We thank K. Williams, K. Stone and A. Barlow for technical help and C. Wunderlich, N. Lindberg and J. Dunne for support. These studies were supported by the NIH and the Howard Hughes Medical Institute.

Subunit of the '20S' proteasome (multicatalytic proteinase) encoded by the major histocompatibility complex

Vianney Ortiz-Navarrete*, A. Seelig†, M. Gernold†, S. Frentzel†, Peter M. Kloetzel† & Günter J. Hämmerling*‡

* Institute for Immunology and Genetics, German Cancer Research Center, Im Neuenheimer Feld 280, 6900 Heidelberg, Germany

† ZMBH/Molecular Genetics, University of Heidelberg, Im Neuenheimer Feld 282, 6900 Heidelberg, Germany

CYTOTOXIC T lymphocytes recognize fragments (peptides) of protein antigens presented by major histocompatibility complex (MHC) class I molecules. In general, the peptides are derived from cytosolic proteins and are then transported to the endoplasmic reticulum where they assemble with the MHC class I heavy chains and β_2 -microglobulin to form stable and functional class I molecules¹⁻⁴. The proteases involved in the generation of these peptides are unknown. One candidate is the proteasome, a non-lysosomal proteinase complex abundantly present in the cytosol. Proteasomes have several proteolytically active sites and are complexes of high relative molecular mass (M_r , about 600K), consisting of about 20-30 subunits with M_r s between 15 and 30K (refs 5-10). Here we show that at least one of these subunits is encoded by the mouse MHC in the region between the *K* locus and the MHC class II region, and inducible by interferon- γ . This raises the

intriguing possibility that the MHC encodes not only the MHC class I molecules themselves but also proteases involved in the formation of MHC-binding peptides.

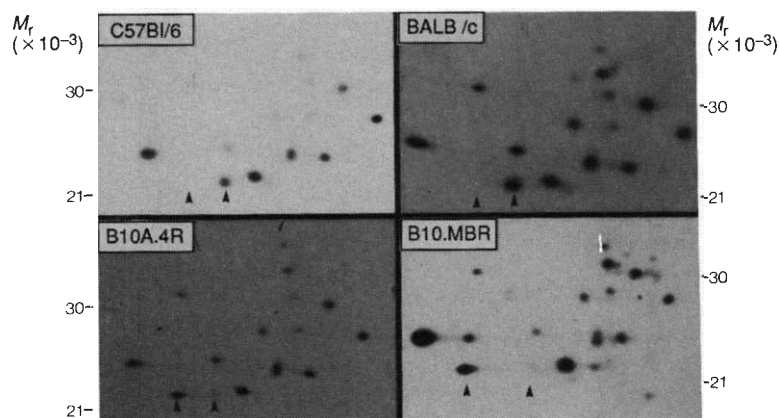
Immunoprecipitation of proteasomes from ³⁵S-methionine-labelled C57Bl/6 spleen cells with a specific antiserum MP1 raised in rabbits against mouse proteasomes¹¹ and subsequent electrophoretic separation of the subunits by two-dimensional gel electrophoresis resulted in a characteristic pattern of about 22-24 protein spots. The precise number of spots was difficult to estimate because some spots seemed to be largely superimposed and others were extremely weak. The respective proteins are known to form the high M_r '20S' proteasome complex¹². To investigate whether different MHC haplotypes would yield different patterns, proteasomes were precipitated from cellular extracts of spleen cells of various mouse strains. Figure 1 shows that precipitation from C57Bl/6(H-2^b) and BALB/c (H-2^d) splenocytes yielded exactly the same two-dimensional gel pattern. Identical results were obtained with splenocytes from BALB.B (H-2^b) and B10.GD (H-2^k) mice. But in H-2^k-derived precipitates one subunit (M_r ~23K) migrated at a different position (arrows in Fig. 1). This suggests that the respective proteasome subunit occurs in allelic forms which are controlled by the MHC. The use of congenic intra H-2 recombinant mice allowed us to assign the gene encoding the allelic protein to the MHC region between the *K* locus and the class II region. The B10.MBR mouse mapped it to the right of the *Pb* locus, and the B10.A(4R) mouse to the left of *Ea*.

The human lymphoblastoid B-cell-derived T2 cell line was obtained by immunoselection with anti-class II antibodies from the parental T1 cell. T2 has a large homozygous deletion encompassing the MHC class II region¹³. This deletion also impaired the supply of intracellular peptides binding to MHC class I molecules and strongly reduced cell-surface expression of the MHC class I (refs 14, 15). Immunoprecipitations of T2 and the

‡ To whom correspondence should be addressed.

FIG. 1 Proteasome pattern correlates with MHC haplotype. ^{35}S -labelled splenocytes from mouse strains with different MHC haplotypes were precipitated with a rabbit anti-mouse proteasome serum and precipitates were separated by two-dimensional gel electrophoresis. Arrows indicate allelic forms of a proteasome subunit with different positions in the *k* haplotype compared with the *b* and *d* haplotypes. The other subunits do not have MHC-linked allelism. With the sequence of the MHC class I and class II loci being *K Pb Ob Ab Aa Eb Ea D* the mouse strains shown carry the following alleles: C57B1/6, (*bbbbb*); BALB/c (*ddddd*); B10.A(4R), (*kkkkkk*) with a crossover between *Eb* and *Ea*; B10.MBR (*bbkkkk*) with a crossover between *Pb* and *Ob* (ref. 32). The strains C3H and B10.BR which have the *k* allele at all loci displayed the same proteasome pattern as B10.A(4R) and B10.MBR.

METHODS. Two-dimensional gel analysis was done as described³³. Briefly, 10^7 splenocytes were cultured for 48 h in Interleukin-2-rich supernatant (10%) from rat spleen cells stimulated for 3 days with concanavalin A and labelled in 1 ml methionine-free RPMI 1640 medium for 6 h with 500 μCi ^{35}S -methionine. After washing the cells were lysed with lysis buffer (2% Nonidet P-40, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM phenylmethylsulphonyl fluoride/aprotinin (30 μg m^{-2}). Incorporation was 5×10^8 c.p.m. Lysates were preabsorbed with 20 μl normal rabbit serum followed by 100 μl *Staphylococcus aureus* Cowan I and centrifugation. Specific precipitation was done overnight with



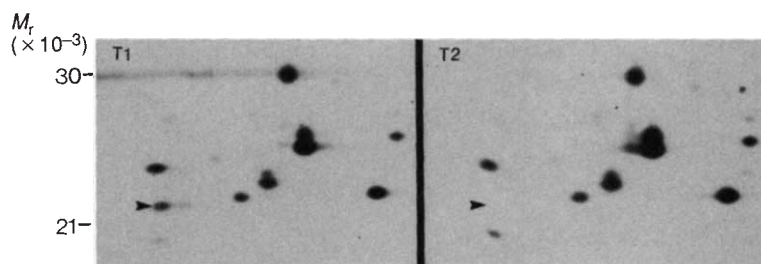
30 μl rabbit anti-mouse proteasome serum MP1 (ref. 11) plus 50 μl protein-A Sepharose. After centrifugation and washing four times, the precipitates were separated using nonequilibrium pH electrophoresis (NEPHGE) for the first dimension (600 V, 6 h) and SDS-PAGE (12%) for the second dimension. Gels were subjected to fluorography and exposed to Kodak XAR-5 film for 4 days.

parental T1 cells were done with the MP1 antiserum which cross reacts with human proteasomes. Indeed, comparison of the two-dimensional gels derived from these cell lines showed that in the case of T2 one subunit ($M_r \sim 23\text{K}$) was missing (fig. 2, arrow). This suggests that the respective protein was controlled by a gene in the MHC class II region which is deleted in T2 cells. Note that this human proteasome subunit is located in the same area on the two-dimensional gel as the allelic forms observed in mice (Fig. 1) indicating that the MHC-linked subunits in mouse and human have similar M_r (23K) and electrophoretic properties and are probably corresponding proteins.

The expression of MHC antigens and the putative peptide transporters encoded by the MHC class II region¹⁶⁻²⁰ can be upregulated by interferon- γ (IFN- γ ; refs 16, 21, 22). When the mouse strain C57B1/6 (*H-2^b*)-derived T lymphoma RMA (ref. 4) was cultured in the presence of IFN- γ the expression of some proteasome subunits was strongly enhanced and others less strongly (Fig. 3). Interestingly, the MHC-encoded subunit belonged to the former group (Fig. 3, arrows). Figure 1 also probably represents enhanced proteasomes because precipitates from splenocytes cultured in the absence of concanavalin A supernatant yielded less intense spots.

If proteasomes are indeed involved in the generation of MHC class I-binding peptides then they might also be expected to control MHC class I cell-surface expression. The mouse strain C57B1/6 (*H-2^b*)-derived lung carcinoma CMT (ref. 23) is characterized by low surface MHC class I expression, owing to the lack of intracellular association of class I heavy chains and β_2 -microglobulin which could be due to insufficient supply with peptides. This cell line also has low levels of proteasomes, and only when the cells are treated with IFN- γ is an increase in MHC class I and proteasome synthesis observed (Fig. 4). But it remains to be determined whether this correlation is meaningful.

FIG. 2 Absence of a proteasome subunit from T2 cells. The parental human T1 cell line and the variant T2 which lacks the HLA class II region were precipitated with the MP1 antiserum which crossreacts with human proteasomes. Note that in T2 cells one subunit is missing as indicated by an arrow. Precipitations were done as described in Fig. 1 with 2.2×10^7 c.p.m. obtained from 10^7 labelled cells. Exposure time of the autoradiographs was 4 days.



We have demonstrated that the MHC controls proteasome subunits both in mouse and in humans. The observed correlation between expression of proteasomes and MHC class I supports the assumption but does not prove that they are involved in antigen processing. According to the two-dimensional gel pattern the proteasomes seem to be very similar or identical to the so-called low M_r proteins (LMP) which could be precipitated with certain polyclonal anti-H-2 sera and some of which are also controlled by the same region of the MHC (refs 24-26). The anti-H-2 sera tested by us, however, did not contain activity against proteasomes. A further similarity is that LMP are also inducible by IFN. The allelic proteasome subunit described here corresponds probably to the LMP-7 protein which is different in *H-2^k* and *H-2^d* haplotypes²⁶, but we did not observe allelic differences of proteasomes between *H-2^b* and *H-2^d* as reported for another LMP protein²⁴⁻²⁶. The possibility of a relationship between LMP and proteasomes and their potential involvement in antigen processing has been pointed out recently²⁷ and is now supported by the present results.

The same region of the MHC also encodes two membrane transporters of the ABC superfamily which may be involved in the transport of peptides to class I molecules¹⁶⁻¹⁹. Consistent with this hypothesis is that transfection of one of these genes into the HLA-defective cell line .134 resulted in normal assembly and cell-surface expression of class I molecules²⁰. The presence of a proteasome gene in the MHC could be fortuitous but it is intriguing to note that the MHC encodes not only the structural genes for the MHC class I heavy chains themselves but probably also the transporters and proteases which are required for antigen processing and assembly of class I molecules with peptides. This would suggest an evolutionary advantage for a genetic linkage of several of the genes necessary for antigen processing and class I expression. During evolution the proteasomes clearly

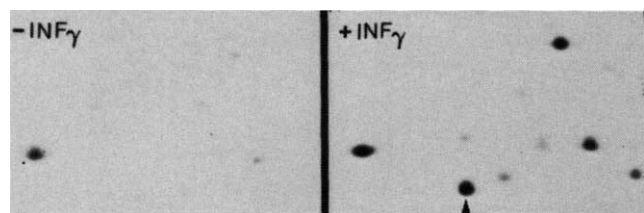


FIG. 3 Enhancement of proteasome expression by IFN- γ . RMA (H-2^b) murine T lymphoma cells⁴ (10^7) were cultured for 48 h without and with 20 units ml⁻¹ rIFN- γ and metabolically labelled. An identical amount of labelled lysate (2.2×10^7 c.p.m.) was used for precipitation to guarantee accurate comparison. Exposure time was 4 days. The MHC-encoded subunit is marked by an arrow. After 15 days exposure all proteasome subunits were also visible in the untreated RMA population.

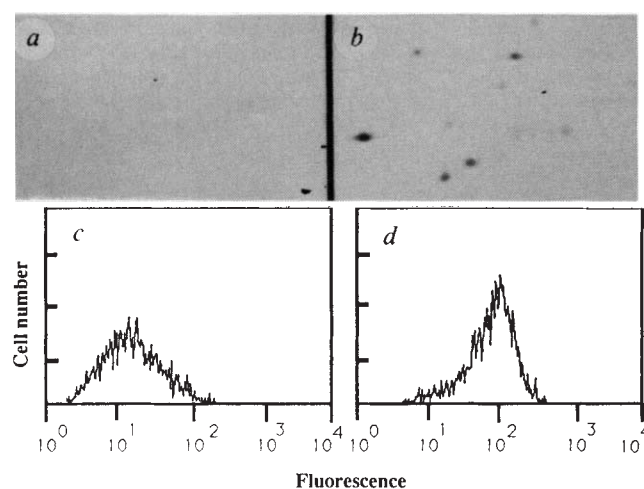


FIG. 4 Low expression of proteasomes in cells with low H-2 class I surface expression. *a, b*, Proteasomes were precipitated from an equivalent number of counts (9×10^6 c.p.m.) obtained from 10^7 metabolically labelled mouse CMT (H-2^b) lung carcinoma cells. *c, d*, Surface staining with antibody B22-249 (D^b) followed by FITC-conjugated goat anti-mouse IgG. CMT cells are low for surface MHC class I molecules but express intracellularly class I and β_2 -microglobulin chains which fail to associate. When this cell line is treated with IFN- γ , it induces assembly and surface class I expression²³. *a, c*, CMT cells cultured without IFN- γ ; *b, d*, CMT cells cultured for 48 h with 20 units ml⁻¹ rIFN- γ . Exposure time of the autoradiographs was 15 days. It seems that CMT cells express less proteasomes than, for example, RMA or T2 cells (Figs 2 and 3), even after IFN treatment.

preceded the immune system as they are found in organisms such as *Archaeobacteria* and *Drosophila*^{9,28-31}. If proteasomes are indeed involved in antigen processing it would seem that the immune system has used preexisting components for the mechanism of antigen processing. From experiments with yeast it has been speculated that proteasomes may be required for the survival of a cell³¹. That a proteasome subunit is missing in the T2 cells suggests that either the respective subunit is not essential for proteasome activity or that there exist multiple and perhaps specialized proteasome complexes in a cell, one of which could be preferentially used for antigen processing. The fact that only some proteasome subunits are inducible by IFN- γ , including the MHC-encoded one, is consistent with the latter assumption. □

- Rivett, J. *Archs Biochem. Biophys.* **268**, 1-8 (1989).
- Kopp, F., Steiner, R., Dahlmann, B., Kuehn, L. & Reinauer, H. *Biochem. biophys. Acta* **872**, 253-260 (1986).
- Tanaka, K., Ichihara, A., Waxman, L. & Goldberg, A. L. *J. biol. Chem.* **263**, 15197-15203 (1988).
- Kloetzel, P. M. *Molec. Biol. Rep.* **12**, 223-227 (1987).
- Kloetzel, P. M. *et al. Biomed. biochim. Acta* **50**, 4-6; 451-457 (1991).
- Falkenburg, P. E. and Kloetzel, P. M. *J. biol. Chem.* **264**, 6660-6666 (1989).
- Seelig, A., Kloetzel, P. M., Kuehn, L. & Dahlmann, B. *Biochem. J.* (in the press).
- Haass, Ch., Pesold-Hurt, B., Multhamp, G., Beyreuther, K. & Kloetzel, P. M. *Gene* **90**, 235-241 (1990).
- DeMars, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8183-8187 (1985).
- Cerundolo, V. *et al. Nature* **345**, 349-352 (1990).
- Hosken, N. A. & Bevan, M. J. *Science* **248**, 367-370 (1990).
- Trowsdale, J. *et al. Nature* **348**, 741-744 (1990).
- Spies, Th. *et al. Nature* **348**, 744-747 (1990).
- Deverson, E. V. *et al. Nature* **348**, 738-741 (1990).
- Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723-1726 (1991).
- Spies, Th. & De Mars, R. *Nature* **351**, 323-325 (1991).
- Lindahl, P., Leary, P. & Gresser, I. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2785-2789 (1973).
- Wallach, D., Fellows, M. & Revel, M. *Nature* **299**, 833-836 (1982).
- Klar, D. & Hämmerling, G. J. *EMBO J.* **8**, 475-481 (1989).
- Monaco, J. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001-3005 (1982).
- Monaco, J. J. & McDevitt, H. O. *Nature* **309**, 797-799 (1984).
- Monaco, J. J. & McDevitt, H. O. *Hum. Immun.* **15**, 416-426 (1986).
- Parham, P. *Nature* **348**, 674-675 (1990).
- Dahlmann, B. *et al. Baumeister, W. FEBS Lett.* **251**, 125-131 (1989).
- Tanaka, K. *et al. J. biol. Chem.* **263**, 16209-16217 (1988).
- Arrigo, A.-P., Simon, M., Darlix, J.-L. & Spahr, P. F. *J. molec. Evol.* **25**, 141-150 (1987).
- Fujiwara, T. *et al. J. biol. Chem.* **265**, 16604-16613 (1990).
- Uematsu, Y., Fischer-Lindahl, K. & Steinmetz, M. *Immunogenetics* **27**, 96-101 (1988).
- Koch, N. & Hämmerling, G. J. *J. Immun.* **128**, 1155-1158 (1982).

ACKNOWLEDGEMENTS. We thank N. Bulbuc for technical help and U. Esslinger for the preparation of the manuscript. This work was partly supported by the Deutsche Forschungsgemeinschaft.

Homology of proteasome subunits to a major histocompatibility complex-linked LMP gene

Coleen K. Martinez & John J. Monaco

Departments of Microbiology & Immunology and Pathology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298-0678, USA

THE class II region of the major histocompatibility complex (MHC) contains genes encoding at least two subunits of a large, intracellular protein complex (the low molecular mass polypeptide, or LMP, complex)¹⁻³. This complex is biochemically similar to the proteasome, an abundant and well conserved protein complex having multiple proteolytic activities. Here we report the isolation of a complementary DNA corresponding to one of the subunits of the LMP complex, LMP-2. The protein predicted from this cDNA sequence closely matches the amino-terminal peptide sequence of a rat proteasome subunit, confirming that the proteasome and the LMP complex share polypeptide subunits. The LMP-2 gene is tightly linked to *HAMI*, a gene thought to be required for translocating peptide fragments of endogenous antigens into the endoplasmic reticulum for association with MHC class I molecules⁴. These observations suggest that the LMP complex may be responsible for generating peptides from cytoplasmic antigen during antigen processing.

Two subunits of the LMP complex (LMP-2 and LMP-7) are polymorphic, and both of these map to the MHC^{1,3}. Our original genetic mapping data indicated that the gene for at least one of these (LMP-2) mapped between the *Pb* and *Ob* class II genes. The left-hand boundary of the region containing this gene, however, was defined by a single recombinant, B10.MBR, and was thus regarded as tentative. Furthermore, initial screening for transcribed sequences between the *Pb* and *Ob* genes led to the identification of seven new genes in this region, none of which seemed to code for LMP subunits⁵. We therefore sought to re-evaluate the original mapping data by analysing new intra-MHC recombinants. Strains R35 and R37 were derived from animals bearing the *Mus musculus castaneus*-derived MHC haplotypes *cas3* (C3) and *cas4* (C4). The crossovers in both

Received 29 July; accepted 18 September 1991.

- Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457-467 (1985).
- Morrison, L. A., Lukacher, A. E., Braciale, B. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903-921 (1986).
- Moore, M., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777-785 (1988).
- Townsend, A. R. M. *et al. Nature* **340**, 443-448 (1989).

strains map in a 37.5-kilobase (kb) DNA segment⁶ telomeric to the *Pb* gene (Fig. 1a). Serological typing shows that the *LMP-2* locus maps to the right of the crossover in both strains (Table 1), confirming that the gene maps to the right (telomeric) of *Pb*.

The crossover in strain B6.CAS3(R23) is in a short DNA segment about 100-kb telomeric to *Pb* (ref. 7). This segment is completely contained in the 26.5-kb DNA segment⁶ containing

the crossover in B10.SBR mice. The *LMP-2* locus maps to the left of the crossover in B10.SBR (ref. 3 and Table 1), but to the right of the *R23* crossover (Table 1). It therefore became important to confirm and refine the location of these crossovers, as a map position for the *LMP-2* locus could be defined only if the B10.SBR crossover lies to the right of that in R23, that is in the order centromere – *R23* – B10.SBR – telomere.

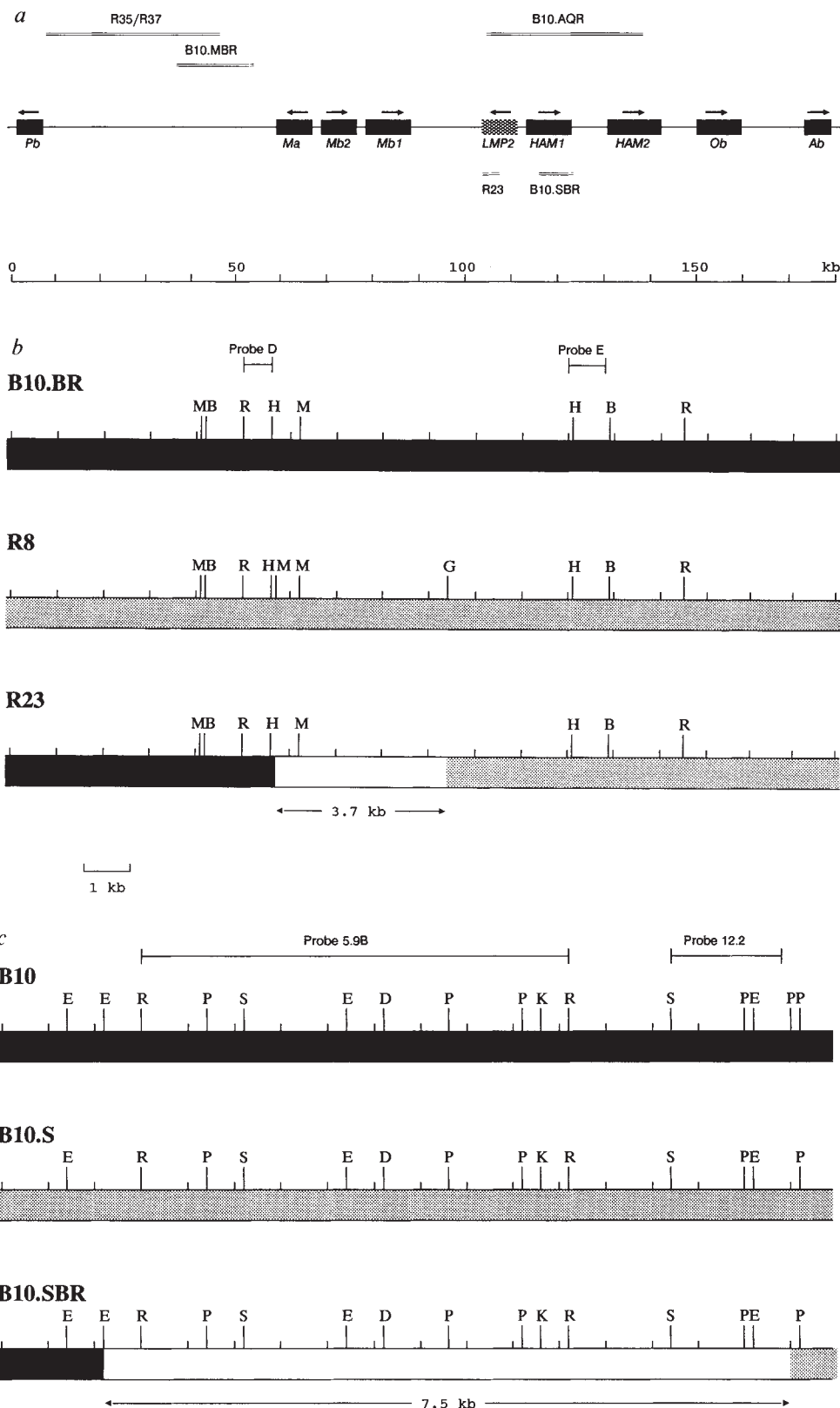


FIG. 1 Crossovers in strains B10.SBR and B6.CAS3(R23) are bracketed by polymorphic restriction sites. **a**, Map of the proximal *H-2* region showing the location of the DNA segments to which the crossovers in the indicated strains are confined, and the position of the putative *LMP-2* gene. **b**, **c**, Partial restriction maps of DNA from parental and recombinant strains. Restriction enzyme abbreviations: B, *Bss*HII; D, *Hind*III; E, *Bst*EI; G, *Bgl*I; H, *Hpa*I; K, *Kpn*I; M, *Msp*I; P, *Pst*I; R, *Eco*RI; S, *Sal*I. Probes used to identify polymorphic restriction sites on Southern blots of genomic DNA are indicated in **b** and **c**. Probe 5.9B is a 4.5-kb *Eco*RI fragment of cosmid 5.9 (see ref. 4), and probe 12.2 is a 1.2-kb *Sal*I fragment of cosmid 5.10 (ref. 6). Probes D and E are derived from digests of cosmid fragment 5.10A, which is the largest of the *Eco*RI fragments of cosmid 5.10; probe D is a 0.6-kb *Hpa*I–*Eco*RI fragment, and probe E is a 0.8-kb *Hpa*I–*Bss*HII fragment. The small (0.9 kb) difference in distance between the polymorphic *Msp*I and *Bgl*I sites between that shown in **b** and the previous report⁶ is probably due to differences in estimating restriction fragment sizes from gels and Southern blots. Boundaries for the genes shown in **a** are approximate^{4,5}. Partial sequence analysis and Southern blots of genomic DNA probed with *LMP-2* cDNA (data not shown) demonstrate that the corresponding gene is composed of at least five exons, is transcribed divergently from the *HAM1* gene, and extends from within 1 kb of the 5' end of *HAM1* at least to within 1 kb of the polymorphic *Msp*I restriction site that defines the leftmost possible position of the *LMP-2* locus.

TABLE 1 LMP types of *H-2* recombinant strains

Strain	K	LMP	A	E
R35	C4	→	C3	C3
R37	C3	→	C4	C4
B10.MBR	b	→	k	k
B10.SBR	b	→	s	s
R23	k	→	C3	C3

Genotypes of individual loci within *H-2* are shown. Vertical bars indicate the positions of the crossover in each strain, and arrows indicate the location of the LMP-2 gene relative to the crossovers. Strains bearing the C4 haplotype contain an allelic form of the LMP complex that is indistinguishable by two-dimensional gel electrophoresis from LMP of the k haplotype, which reacts with murine anti-LMP alloantisera. The sera used in these studies were BALB.B anti-BALB/c and (BALB.B × B10.AQR)_{F1} anti-B10.MBR. Strains bearing the C3 haplotype do not contain LMP complexes reactive with these sera. Both sera presumably recognize an antigenic determinant on LMP subunit number 2 (ref. 3), and hence, the 'lmp locus' being mapped using these sera probably represents the structural gene for the LMP-2 polypeptide. All strains were typed both by direct immunoprecipitation and by cold competition, as previously described^{1,3}.

Southern blots of genomic DNA from B6.CAS3(R23) and from strains bearing the parental haplotypes, B10.CAS3(R8) and B10.BR, were used to confirm previously identified *Msp*I and *Bgl*II polymorphisms⁶, and to define an interval of 3.7 kb to which the R23 crossover is confined (Fig. 1b). The right-hand boundary of the interval containing the B10.SBR crossover was confirmed using probe 12.2 (ref. 6), which detects a polymorphic *Pst*I site (Fig. 1c). The left-hand boundary of this interval is defined by a polymorphic *Bst*EII site, confining the B10.SBR crossover to a 7.5-kb region between these *Bst*EII and *Pst*I restriction sites (Fig. 1c). Thus, the B10.SBR crossover indeed maps to the right of the R23 crossover, and these two crossovers define a region of from 10 to 21 kb that must contain the polymorphic lmp locus (Fig. 1a). This interval contains the entire *HAMI* gene (about 11 kb; M. Attaya and J.M.M., unpublished observations), and about 2 kb of DNA telomeric and 8 kb of DNA centromeric to this gene. Genomic probes from the 8-kb region were found to hybridize to a single 1.3-kb transcript on northern blots of WEHI-3 (*H-2*^d, myelomonocytic cell line) RNA (not shown), consistent with the size range of LMP transcripts³. These probes were used to screen a WEHI-3 cDNA library⁵. The sequence of an 806-base pair (bp) cDNA clone (clone W10) isolated in this manner (Fig. 2a) contains a single long open reading frame. Translation of the open reading frame from the first methionine codon at position 22 to the termination codon at position 679 predicts a 219-amino-acid protein with relative molecular mass (*M*_r) of 23,396 and pI of 4.96; these values agree reasonably well with those determined for the polymorphic subunit number 2 (LMP-2) of the LMP complex¹⁻³ (but see below), consistent with the possibility that the W10 clone in fact corresponds to the LMP-2 gene. Furthermore, anti-LMP antisera apparently are specific for an epitope on the LMP-2 subunit of the *H-2*^d, but not the *H-2*^b allele³. Expression of the W10 (*H-2*^d) cDNA clone in *H-2*^b type cells (data not shown) results in expression of immunoprecipitable (*H-2*^d type) LMP complex in these cells. Taken together, these data argue strongly that the W10 clone derives from the LMP-2 gene.

Brown *et al.*⁸ demonstrated that the LMP complex is closely related to an evolutionarily well conserved protein complex (the proteasome) that possesses multiple proteolytic activities. A search of the National Biomedical Research Foundation protein sequence database using the program FASTA⁹ revealed a striking match between amino acids 21 to 41 of the predicted LMP-2 sequence and the published amino-terminal peptide sequences of purified rat¹⁰ and human¹¹ proteasome subunits (Fig. 2b). If

the published peptide sequence is the actual amino-terminus of the mature protein, then the 20 amino acids preceding this sequence predicted from the cDNA clone must be removed post-translationally. Alternatively, the published sequence may reflect partial proteolysis of this subunit during purification, or may derive from a gene related to, but distinct from, LMP-2. We cannot at present distinguish between these alternatives, although the size of the predicted LMP-2 protein, assuming the published peptide sequence is the amino terminus (*M*_r = 21,325, pI = 4.77), is closer to that estimated from two-dimensional gels.

a

1	GCCGAGCCCGCTCTGCTGAGATGCTGCGGGCAGGAGCACCTACCGCCGG	10
	M L R A G A P T A G	
51	CTCGTTCCGGACGGAAGAAGTCCACACCGGGACAACCATCATGGCAGTGG	26
	S F R T E E V H T G T T I M A V	
101	AGTTTGACGGGGTGTCTGGTGGGCTCTGATTCCCGGGTGTGACGAGGA	43
	E F D G G V V V G S D S R V S A G	
151	ACAGCAGTGGTGAACCGCGTGTTCGACAAGCTCTCCCTCTGCACACGCG	60
	T A V V N R V F D K L S P L H Q R	
201	CATCTTCTGTGCCCTCTCGGGTTCGGCTGCTGATGCCAAGCCATAGCTG	76
	I F C A L S G S A A D A Q A I A	
251	ACATGGCCCGCTACAGCTGGAGCTACACGGGTGGAGCTGGAGAGCCA	93
	D M A A Y Q L E L H G L E L E E P	
301	CCCCTCGTTCTGGCTGCTGCAAACTGGTGAAGAATCTCTCTACAAGTA	110
	P L V L A A A N V V K N I S Y K Y	
351	CCGTGAGGACTTGTAGCGCATCTCATAGTAGCTGGGTGGGACCAACGTG	126
	R E D L L A H L I V A G W D Q R	
401	AGGGGGACAGGTGTATGGAACCTGGGAGGGATGCTAATTCGACAGCCC	143
	E G G Q V Y G T M G S A A D A Q A I A	
451	TTTACCATCGGCGGTTCTGGAAGCTCCTACATTTACGGTTATGTGGACGC	160
	F T I G G S G S S Y I Y G Y V D A	
501	AGCTTATAAGCCAGGCATGACCCCTGAGGAGTGCCGCGGTTTACCACAA	176
	A Y K P G M T P E E C R R F T T	
551	ATGCCATCACTCTGGCCATGAACCGAGATGGCTAGTGGGGGTGTCTATC	193
	N A I T L A M N R D G S S G G V I	
601	TACCTGGTCACCATCACAGCTGCTGGTGGACCATCGAGTCATCTGGG	210
	Y L V T I T A A G V D H R V I L G	
651	AGATGAGCTGCCAAAATTCTACGATGAGTGACTGATCCCCAGAAGTCCCT	219
	D E L P K F Y D E *	
701	TCCTTGTGTGTAATAAACTTTCTGGAACAGAGGCTGGTGCATGGGCA	
751	AAGGTGAAATATGTGTATCAGAGACACGGTGTGTGGAGGGAAGCGTCC	
801	TCCCGG	

b

Rat chain 7	TTIMAVEFDGGVVVGSDSDVS
LMP-2	TTIMAVEFDGGVVVGSDSRVS
Human δ chain	XXIMAVQFDGGVVVGADSRRT

FIG. 2 Nucleotide and deduced amino-acid sequence (single-letter code) of the LMP-2 gene. a. The 806-bp cDNA clone W10 contains a single open reading frame that encodes a 219 amino-acid protein. The region corresponding to amino-terminal amino-acid sequence of purified proteasome subunits is shaded. The polyadenylation signal is underlined and the stop codon identified with an asterisk. b. Comparison of the deduced amino-acid sequence of LMP-2, and the actual N-terminal sequence of rat proteasome chain 7 (ref. 10), and the human macropain (proteasome) δ chain¹¹. The rat and human sequences have 95% and 74% amino-acid identity, respectively, to the corresponding portion of the LMP-2 sequence. Three fragments (1.0-kb *Apal*-*Hind*III, 1.3-kb *Apal*-*Hind*III, and 2.3-kb *Apal*-*Eco*RI) purified from a 9.6-kb *Eco*RI fragment (5.10A) of cosmid 5.10 were used to screen a WEHI-3 myelomonocytic cDNA library⁵. The 806-bp *Eco*RI insert from clone W10 was subcloned into plasmid pGEM-3Zf(+) for sequencing. This sequence will appear in the GenBank and EMBL nucleotide sequence databases.

No sequence motifs characteristic of known proteinases are found in the LMP-2 sequence using the program MOTIFS⁹ and the Prosite database, and none of the five previously isolated cDNA clones for rat proteasome subunits¹²⁻¹⁶ is significantly related to the LMP-2 sequence. Preliminary results suggest that none of those five cDNAs are encoded between *H-2K* and *HAM2* (M. G. Brown and J.J.M., unpublished observations). Thus, the LMP complex may contain both MHC-linked and non-MHC-linked subunits.

The position of the *LMP-2* gene in relation to the *HAM* transporter genes is analogous to that of the substrate-binding protein gene in many of the homologous bacterial transport systems. Thus, the LMP-2 protein may possess a similar function, that is to couple the proteasome and its output (peptides) to the HAM transporter. The unusual structure of this proteolytic complex may be important for generating peptides able to bind class I molecules.⁸ □

Received 30 July; accepted 20 September 1991.

1. Monaco, J. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001-3005 (1982).
2. Monaco, J. J. & McDevitt, H. O. *Nature* **309**, 797-799 (1984).
3. Monaco, J. J. & McDevitt, H. O. *Hum. Immun.* **15**, 416-426 (1986).
4. Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723-1726 (1990).
5. Cho, S., Attaya, M., Brown, M. G. & Monaco, J. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5197-5201 (1991).
6. Uematsu, Y., Fischer-Lindahl, K. & Steinmetz, M. *Immunogenetics* **27**, 96-101 (1988).
7. Uematsu, Y., Kiefer, H., Schulze, R., Fischer-Lindahl, K. & Steinmetz, M. *EMBO J.* **5**, 2123-2129 (1986).
8. Brown, M. G., Driscoll, J. & Monaco, J. J. *Nature* **353**, 355-357 (1991).
9. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387-395 (1984).
10. Lilley, K., Davison, M. & Rivett, A. J. *FEBS Lett.* **262**, 327-329 (1990).
11. Lee, L. et al. *Biochim. biophys. Acta* **1037**, 178-185 (1990).
12. Fujiwara, T. et al. *Biochemistry* **28**, 7332-7340 (1989).
13. Tanaka, K. et al. *Biochemistry* **29**, 3777-3785 (1990).
14. Tamura, T. et al. *FEBS Lett.* **264**, 91-94 (1990).
15. Tanaka, K. et al. *Biochem. biophys. Res. Commun.* **171**, 676-683 (1990).
16. Kumatori, A. et al. *FEBS Lett.* **264**, 279-282 (1990).

ACKNOWLEDGEMENTS. We thank K. Fischer-Lindahl for inbred and recombinant mice bearing *Mus musculus castaneus*-derived MHC haplotypes, and T. Reynolds and the Massey Cancer Center Molecular Biology Core Facility for oligonucleotide synthesis and aid in DNA sequencing. This work was supported by the National Institute of General Medical Sciences (to J.J.M.).

Second proteasome-related gene in the human MHC class II region

Adrian Kelly, Stephen H. Powis, Richard Glynne, Elizabeth Radley, Stephan Beck & John Trowsdale

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

ANTIGEN processing involves the generation of peptides from cytosolic proteins and their transport into the endoplasmic reticulum where they associate with major histocompatibility complex (MHC) class I molecules¹⁻⁶. Two genes have been identified in the MHC class II region, *RING4* and *RING11* in humans, which are believed to encode the peptide transport proteins⁷⁻¹². Attention is now focused on how the transporters are provided with peptides. The proteasome, a large complex of subunits with multiple proteolytic activities, is a candidate for this function¹³⁻¹⁶. Recently we reported a proteasome-related sequence, *RING10*, mapping between the transporter genes¹⁷. Here we describe a second human proteasome-like gene, *RING12*, immediately centromeric of the *RING4* locus. Therefore *RING12*, 4, 10 and 11 form a tightly linked cluster of interferon-inducible genes within the MHC with an essential role in antigen processing.

The *RING12* clone was isolated by screening a B-cell complementary DNA library with fragments of cosmid U15¹⁸. *RING12* mapped within 500 base pairs (bp) of *RING4*, with the promoter (5') ends of the two genes adjacent. Northern blot analysis detected a 900-bp messenger RNA transcript in both B and T lymphocytes (Fig. 1). Expression of mRNA was strongly

upregulated by γ -interferon in epithelial and fibroblast cell lines, although basal levels were low.

The cDNA insert of the *RING12* clone contained a single long open reading frame encoding a 219-amino-acid product (Fig. 2a). The derived protein was similar to fragments of N-terminal amino acid sequence from rat and human proteasome components and appeared most like rat chain 7 (19 out of 21 identical amino acids; Fig. 2b)^{19,20}. The similarity between proteasome and a murine complex of low-molecular-weight polypeptides (LMPs) has already been noted, with the implication that the two are related^{13,14,16}. Two polymorphic LMPs were originally mapped to the mouse class II region and are probably the equivalents of human *RING12* and *RING10*²¹.

The discovery of a second proteasome-like gene in the class II region of the MHC provides further evidence that proteasomes, or related products, are involved in antigen processing. Proteasome components are expressed constitutively and make up a major fraction of total cell protein¹⁵. As *RING10* and *RING12* are expressed at a very low level in some cells, but are inducible with γ -interferon, they may be specifically tailored to antigen presentation and could help recruit housekeeping proteasome components for this function.

Class I and class II MHC antigens are highly polymorphic, and it is of interest to determine the degree of variation within the putative antigen-processing genes. We have already reported that *RING11* is polymorphic and, in characterizing *RING12*, have observed a single amino-acid difference between the cDNA and a genomic sequence (His-Arg, position 60; Fig. 2a)¹². The clustering of polymorphic genes within the MHC may provide an evolutionary advantage by ensuring that particular alleles are maintained in combination²². We are examining the extent

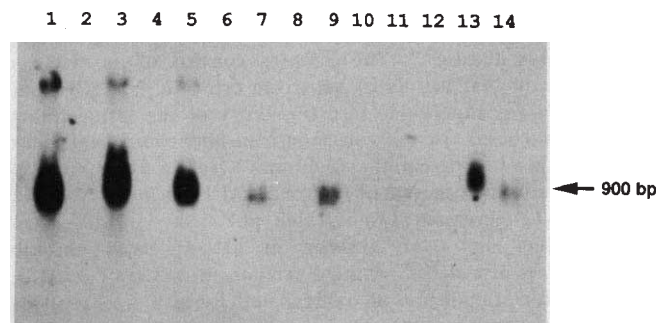


FIG. 1 Expression of *RING12* mRNA on northern blots. The 900-bp band corresponds to the *RING12* transcript and the larger band is from a previous hybridization with the 2.5-kilobase (kb) *RING11* cDNA. Colon carcinoma (tracks 1-6), fibroblast cell line samples (tracks 7-12), Molt4 T cell (track 13) and Raji B cell line (track 14), were loaded (15 μ g total RNA) as follows: tracks 1 and 2, CC20; 3 and 4, SW620; 5 and 6, SW1222; 7 and 8, PAF; 9 and 10, SV80; 11 and 12, MRC5SV. Cells in odd lanes (1 to 11) were treated for 48 h with 250 units ml^{-1} γ -interferon. Total mRNA was isolated by lysis in guanidinium isothiocyanate and purified by centrifugation in caesium chloride solution²⁷.

a

```

      M L R A G A P T G D L P R A G      15
CGGTGCTTGCAGGATGCTGCGGGCGGGAGACCAACCGGGGACTTACCCCGGGCGGGA 60

E V H T G T T I M A V E F D G G V V M G      35
GAAGTCCACACCGGGACCACTCATGGCAGTGGAGTTTGACGGGGCGTGTGTATGGGT 120

S D S R V S A G E A V V N R V F D K L S      55
TCTGATTCGCCAGTGTCTGCAGCGAGGCGTGGTGAACCGAGTGTGGACAGCTGTC 180
R

P L H E H I Y C A L S G S A A D A Q A V      75
CCGCTGCACGAGCAGTCTACTGTGCACTCTCTGGTTACAGTGTGTATGCCAAGCGGTG 240
cgc

A D M A A Y Q L E L H G I E L E E P P L      95
GCTGACATGGCGCCCTACCAGCTGGAGCTCCATGGATAGAAGTGGAGGAACCTCCACTT 300

V L A A A N V V R N I S Y K Y R E D L S      115
GTTTGGCTGCTGCAATGTGGTGAAGAAATATAGCTATAAATATCAGAGGAGCTTGTCT 360

A H L M V A G W D Q R E G G Q V Y G T L      135
GCACATCTCATGGTAGCTGGTGGGACCAACGCGAAGAGGTGAGGTATATGAACCCCTG 420

G G M L T R Q P F A I G G S G S T F I Y      155
GGAGGAATGCTGACAGCAGCTTTTGCCATTGGCTGCGGACGACCTTTATCTAT 480

G Y V D A A Y K P G M S P E E C R R R F T      175
GGTATGTGGATGCAGCATATAAGCAGGAGTGTCTCCCGAGGAGTGCAGGCGCTTACC 540

T D A I A L A M S R D G S S G G V I Y L      195
ACAGAGCTGATTGCTTGGCCATGAGCCGGAGTGGCTCAAGAGGGGGTGTCTATCTG 600

V T I T A A G V D H R V I L G N E L P K      215
GTCATATTACAGCTGCCGGTGTGACCATCGAGTCATCTTGGGCAATGAAGTGCACAAA 660

F Y D E *      219
TTCTATGATGAGTGAACCTTCCCGAGCTTCTCTTTCTATTGTAATAAACTC      715

```

b

	amino acid 19	amino acid 46
RING12	TGTTIMAVEFDDGGVVMGSDSRVSAGEAV	
RAT CHAIN 7	-----V-----D--	
RAT CHAIN 5	-----Q-----L-A--	
HUMAN DELTA	-----Q-----L-A---TTT-SYI	
RING10	A---TL-FK-QH---IAAV---A---SYI	
RAT CHAIN 6	---L-FK-QE---ILAX---	
HUMAN EPSILON	L-FK-RX---IVAA---AT---GYF	
HUMAN BETA	---SVLG-K-E---IAA-MLG	
RAT CHAIN 1	---AG-VYKD-I	
HUMAN ALPHA	-AG-VYKD-I-L-A-TXATXVI	
RAT CHAIN 2	---AGLV-KD---IL-A-XEATN	
RAT CHAIN 3	-QNPM-TGTSV---AKF	
RAT CHAIN 4	SFSPYA-N---TLLVIXE	
HUMAN GAMMA	FSPYV-N---TILAIAGEDF-IV-	

FIG. 2 *a*, Sequence of the *RING12* cDNA clone. The derived amino-acid sequence of the longest open reading frame is shown in single-letter code above the sequence. The 715-bp insert encodes a protein of 219 amino acids with a predicted M_r of 23.2K and a pI of 4.64. The TGA stop codon (*) was followed by a 33-bp 3' untranslated region containing a AATAAA poly(A) addition sequence (underlined). There is a potential N-linked glycosylation site (105) and a potential disulphide bridge at the two cysteine residues (63 and 171). When compared with the full genomic sequence (from cosmid U15)¹⁸ a single base-pair difference was observed (G for A (194), His for Arg (60)). The sequence shown is probably full-length: the cDNA clone is about the same size as products detected on northern blots and the 5' genomic sequence reveals no upstream AUG codons before reaching the adjacent gene *RING4* (data not shown). In addition, although the predicted initiating AUG codon conforms only poorly to the eukaryotic translation initiation consensus²⁸, the sequence about 18 residues downstream of this Met codon matches limited N-terminal peptide data from related proteins (Fig. 2*b*). The *RING12* and *RING10* amino-acid sequences were closer to each other (30% identity) than to human proteasome cDNA sequences (17–23%; ref. 25, data not shown, but see Fig. 2*b*), suggesting that they arose by duplication of an initial transporter and proteasome gene pair with subsequent inversion of one of the genes. The close association of the *RING12/4/10/11* genes may effect coordinated regulation: all four genes are similarly expressed and the promoters of the *RING12* and *RING4* genes share a potential interferon response element (data not shown)²⁶. High-stringency hybridization of the cDNA probe to human genomic DNA on Southern blots showed that *RING12* is a single-copy gene. Unlike *RING10*, *RING12* showed no strong homology to protease active site motifs in the PROSITE database^{17,29}. It is possible, however, to align sequences at amino acid 147–151(GGSGS) and 187–191(GSSGG), assuming one mismatch, with some serine protease active sites²⁹. The NBC cDNA library was prepared in the CDM8 vector as described^{30,31}. DNA sequencing was performed using the random dideoxynucleotide approach³², and sequences were assembled using the SAP program³³. *b*, Comparison of the *RING12* amino acid sequence with those of other proteasome components. Protein sequences were screened against Swiss Prot 17 and PIR27 databases. Only partial amino acid sequence data was available for these proteasome components. The sequences, aligned to *RING12* amino acid 19–46, were from references as follows: rat chains 1–7 (ref. 19), human chains α – ϵ (ref. 20) and *RING10* (ref. 17). Identical residues are marked – X, unknown residue. The proposed *RING12* sequence was at least 18 amino acids longer at the N-terminus than published peptide fragment sequences. This may indicate post-translational processing of the *RING12* protein.

of linkage disequilibrium between the *RING* genes and class I and class II alleles in normal populations and those with MHC-associated disease^{23,24}. □

Received 17 September; accepted 2 October 1991.

1. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457–467 (1985).
2. Townsend, A. & Bodmer, H. A. *Rev. Immun.* **7**, 601–24 (1989).
3. Moore, M., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777–785 (1988).
4. Yewdell, J. W. & Bennink, J. R. *Cell* **62**, 203–206 (1990).
5. Townsend, A. et al. *Nature* **340**, 443–8 (1989).
6. Morrison, L. A., Lukacher, A. E., Braciale, B. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903–921 (1986).
7. Deverson, E. V. et al. *Nature* **348**, 738–41 (1990).
8. Monaco, J. J., Cho, S. & Attaya, M. *Science* **250**, 1723–1726 (1990).
9. Spies, T. et al. *Nature* **348**, 744–7 (1990).
10. Spies, T. & DeMars, R. *Nature* **351**, 323–324 (1991).
11. Trowsdale, J. et al. *Nature* **348**, 741–4 (1990).
12. Powis, S. H. et al. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
13. Monaco, J. J. & McDevitt, H. O. *Nature* **309**, 797–799 (1984).
14. Monaco, J. J. & McDevitt, H. O. *Hum. Immun.* **15**, 416–26 (1986).
15. Rivett, A. J. *Arch. Biochem. Biophys.* **268**, 1–8 (1989).
16. Parham, P. *Nature* **348**, 674–5 (1990).
17. Glynn, R. et al. *Nature* **353**, 357–360 (1991).
18. Blanck, G. & Strominger, J. L. *J. Immun.* **141**, 1734–7 (1988).
19. Lilley, K. S., Davison, M. D. & Rivett, A. J. *FEBS Lett.* **262**, 327–329 (1990).
20. Lee, L. W. et al. *Biochim. biophys. Acta* **1037**, 178–185 (1990).
21. Monaco, J. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3001–3005 (1982).
22. Lawlor, D. A., Zemmour, J., Ennis, P. D. & Parham, P. A. *Rev. Immun.* **8**, 23–63 (1990).
23. Dupont, B. *Immunobiology of HLA* (Springer, New York, 1989).
24. Tiwari, J. L. & Terasaki, P. I. *HLA and Disease* (Springer, New York, 1985).
25. Tamura, T. et al. *Biochim. biophys. Acta* **1098**, 95–102 (1991).
26. Friedman, R. L. & Stark, G. R. *Nature* **314**, 637–639 (1985).
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Press, New York, 1982).
28. Kozak, M. *Cell* **44**, 283–292 (1986).
29. Barroch, A. *Prosite: A Dictionary of Protein Sites and Patterns* 5th edn (Département de Biochimie Médicale, Université de Genève, Geneva, 1990).
30. Seed, B. *Nature* **329**, 840–842 (1987).
31. Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8573–8577 (1987).
32. Bankier, A. T., Weston, K. M. & Barrell, B. G. *Meth. Enzym.* **155**, 51–93 (1987).
33. Staden, R. *CABIOS* **6**, 387–393 (1990).

ACKNOWLEDGEMENTS. Cosmid clone U15 was a kind gift from George Blanck, S.H.P. is a MRC Clinician Scientist Fellow. We thank J. Monaco and G. Hammerling for unpublished information and A. Townsend for discussions.

Activation of the NADPH oxidase involves the small GTP-binding protein p21^{rac1}

Arie Abo*, Edgar Pick*, Alan Hall†, Nicholas Totty‡, Carmel G. Teahan§ & Anthony W. Segal§

* Laboratory of Immunopharmacology, Sackler School of Medicine, Tel Aviv University, Israel

† Chester Beatty Laboratories, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

‡ Ludwig Institute for Cancer Research, 91 Riding House Street, London W1P 8BT, UK

§ Department of Medicine, University College London, Rayne Institute, University Street, London WC1E 6JJ, UK

PROFESSIONAL phagocytes, such as neutrophils and monocytes, have an NADPH oxidase that generates superoxide and other reduced oxygen species important in killing microorganisms (reviewed in ref. 1). Several components of the oxidase complex have been identified as targets of genetic defects causing chronic granulomatous disease^{2–4}. The complex consists of an electron transport chain that has as its substrate cytosolic NADPH and which discharges superoxide into the cavity of the intracellular phagocytic vacuole. The only electron transport component identified so far is a low-potential cytochrome *b* (refs 5, 6), apparently the only membrane component required⁷. At least three cytosolic factors are also necessary, two of which, p67^{phox} and p47^{phox}, have been identified by their absence in patients with chronic granulomatous disease^{8–11}. A third component, σ 1 (refs 12, 13), is required for stimulation of oxidase activity in a cell-free system^{14–16}. The active components of purified σ 1 are two proteins

that associate as heterodimers¹⁷, and here we report that these are the small GTP-binding protein p21^{rac1} and the GDP-dissociation inhibitor rhoGDI.

We have sequenced the two component proteins of $\sigma 1$, p22 and p26 (relative molecular masses 22,000 and 26,000 respectively). The amino termini of both proteins were found to be blocked on direct sequencing after separation by SDS-polyacrylamide gel electrophoresis and blotting onto Immobilon membrane. In view of the small quantities of protein available, and the fact that they copurify, we analysed peptides separated by reverse-phase chromatography from a tryptic digest of a mixture of p22 and p26 (Fig. 1, lane 1). Of the thirteen sequences obtained, seven were homologous with four regions of p21^{rac1} (Fig. 2), which is a small GTP-binding protein initially identified in myeloid cells¹⁸ and sequenced from H160 cells. Another three of the sequences were identical with those of three regions of bovine brain GDP-dissociation inhibitor¹⁹ (GDI), a regulatory protein that binds to the GTP-binding protein p21^{rho}. There was no homology with the remaining three sequences, GYPE-VGLYK, XVLVLLMV and XVEVLKEDVM (one-letter code; Swiss protein database), which were present at concentrations roughly equimolar with the p21^{rac1} and GDI sequences. They could either be derived from non-homologous regions of p22 or p26 or from a contaminating protein, although only two main protein bands were observed in Coomassie blue- and silver-stained SDS-polyacrylamide gels.

In view of the different sizes of p21^{rac1} and GDI, the smaller protein was more likely to be p21^{rac1}. To test this we used a commercial antibody against myosin light chain from chicken lens which crossreacts with p22 of $\sigma 1$ (A.A. and E.P., unpublished result; Fig. 1, lane 2). Figure 1, lane 3, shows that the antibody also recognizes recombinant p21^{rac1} on western blots.

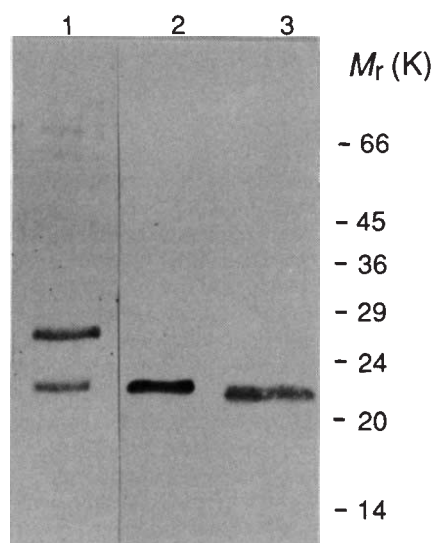


FIG. 1 Sigma 1 consists of two subunits with apparent relative molecular masses (M_r) of 22,000 and 26,000 (22K and 26K) on SDS/PAGE as shown by Coomassie blue staining (lane 1). The 22K component (lane 2) and recombinant p21^{rac1} (lane 3) both react on western blots with an antibody to myosin light chain. No reaction was seen with recombinant Ha-ras or rhoA proteins (not shown).

METHODS. Sigma 1 was purified from cytosol separated from guinea pig macrophages harvested 6 days after the intraperitoneal injection of oil³⁶. Purification was achieved by ammonium sulphate fractionation, hydrophobic chromatography on phenyl-Superose, absorption by CM-Sephacrose, ion-exchange chromatography on DEAE-Sephacrose and gel filtration on Superose 12 (A.A. and E.P., manuscript submitted). p21^{rac1} was partially purified from an *Escherichia coli* expression system as described³⁷. Both this protein (200 ng) and $\sigma 1$ (200 ng of each protein) were analysed on western blots developed for 1 h with a 1:400 dilution of antibody to chicken lens myosin light chain (Sigma), followed by an alkaline phosphatase conjugate of goat antimouse IgG (heavy and light chains, 1:5,000 for 1 h; Jackson Laboratories).

With the protein sequence information, we conclude that p22 is p21^{rac1} and p26 is rhoGDI.

To analyse the possible function of p21^{rac1} in activating the generation of superoxide, we attempted to replace the $\sigma 1$ component in a cell-free system with recombinant p21^{rac1}. Figure 3a shows that 300 ng of partially purified p21^{rac1} increases superoxide production by threefold. Furthermore, activation only occurs when p21^{rac1} is first preincubated with GTP- γ -S. A similar dose response is observed between the concentration of $\sigma 1$ or p21^{rac1} and superoxide generation (Fig. 3b). Neither p21^{rap1} in the GDP form, nor GTP- γ -S alone gives measurable stimulation. Recombinant ras or rho proteins were inactive.

We conclude that the active component of $\sigma 1$ is the small GTP-binding protein, p21^{rac1}. This observation is consistent with a considerable body of evidence in support of the involvement of a GTP-binding protein in the cell-free system²⁰⁻²⁴. There is an association between the cytochrome *b*₂₄₅ of this oxidase system and a membrane-bound small GTP-binding protein, p21^{rac1}, with which it copurifies and cross-immunoprecipitates²⁵. The proposed involvement of p21^{rap1} in this oxidase was purely dependent on its association with the cytochrome *b* and did not include functional studies as demonstrated here for the cytosolic p21^{rac1}. We have found a progressive and eventually complete dissociation of p21^{rap1} as the cytochrome is purified (A.W.S., unpublished result); in experiments in which cytochrome *b* is used as the only membrane component in a fully functional cell-free system, p21^{rap1} is undetectable in the cytochrome preparation (E.P., unpublished result). Whether or not this small GTP-binding protein also plays a part in the oxidase remains to be established.

It is interesting that $\sigma 1$, which was purified because of its ability to activate the NADPH oxidase, is apparently active in the absence of added GTP, whereas recombinant p21^{rac1} needs preincubation with GTP- γ -S. This discrepancy can be explained if a guanine nucleotide-exchange factor for p21^{rac1}, along with cell-derived GTP, are present in the cytosolic fraction added to the assay. In this way the complex of GDI with p21^{rac1}·GDP would be converted into uncomplexed p21^{rac1}·GTP. As nucleotide-exchange factors for rho-like proteins act only on post-translationally processed protein, the recombinant p21^{rac1} would not be similarly activated²⁶. p20^{rhoB} binds to membranes and the GDP-bound form associates with GDI and is displaced from the membranes by it²⁷. GDI might provide a mechanism for recycling p21^{rac1}. GDP between the membranes and cytosol in which it is susceptible to nucleotide exchange.

The mechanism by which p21^{rac1} contributes to activation of the oxidase system is not clear. It belongs to the rho subfamily of related ras proteins and the prototype for this family, p21^{rho}, regulates actin polymerization in various cells^{28,29}. n47^{phox} and

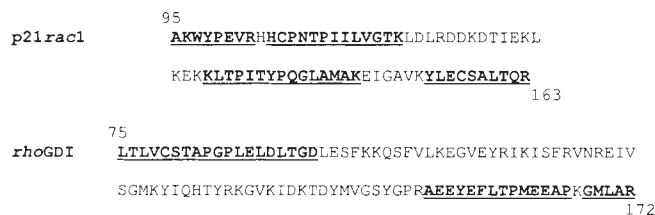


FIG. 2 Sequence homology between peptides obtained from a tryptic digest of $\sigma 1$ containing a mixture of the 22K and 26K proteins and p21^{rac1} and rhoGDI respectively. Sequences of the peptides are shown in bold type and underlined, and are identical to sequences of the proteins, the relevant portions of which are shown. p21^{rac1} was identified as p22 component by western blot analysis (see Fig. 1).

METHODS. A mixture (20 μ g protein) of the 22K and 26K proteins was digested with trypsin, the peptides separated by reverse-phase liquid chromatography and sequenced on an Applied Biosystems Model 477A Pulse Liquid sequencer. Yields of PTH-amino acids at the first step of each sequence run varied from 6–25 pmol. The three non-homologous peptides (see text) yielded 7, 10 and 10 pmol.

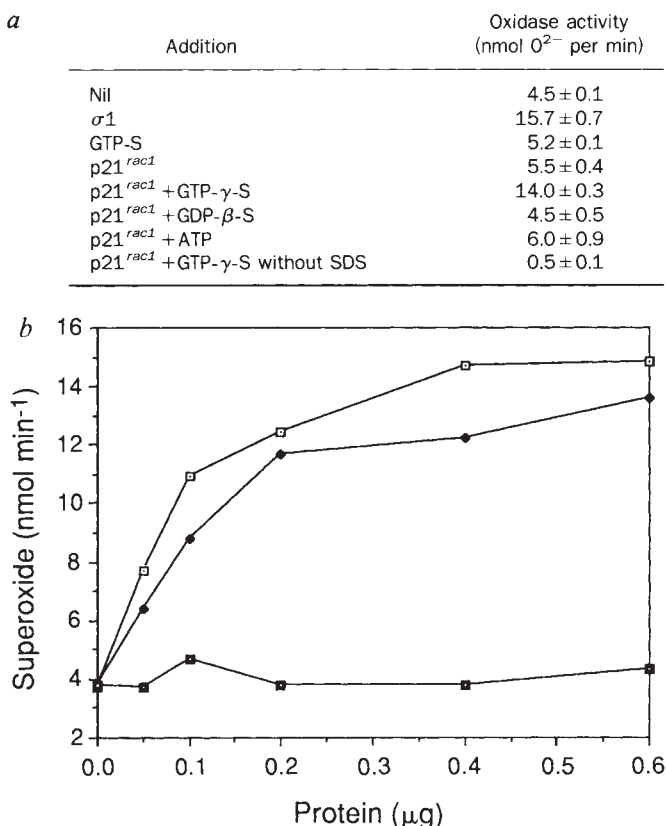


FIG. 3 Effect of recombinant p21^{rac1} alone, or after nucleotide exchange, on the activity of the NADPH oxidase in the cell-free system depleted of σ1 activity. **a**, Collated results of 3 to 6 experiments, and **b**, Dose response of σ1 □, and GTP-γ-S ♦ and GDP-β-S ■ forms of p21^{rac1}. **METHODS.** The cell free system contained macrophage membranes (12 μg protein per assay solubilized with octyl glucoside³⁸) and the 37% ammonium sulphate precipitate of macrophage cytosol (42 μg protein) which is depleted of σ1 activity but contains p47^{phox} and p67^{phox}. The reaction was measured 90 s after the addition of an optimal concentration of SDS (80–120 μM) by superoxide dismutase-inhibitable cytochrome *c* reduction³⁸. For data presented in **a**, pure σ1 (200 ng¹⁷) or partially purified recombinant p21^{rac1} (300 ng) was added. Recombinant p21^{rac1} is predominantly in the GDP-form after purification from *E. coli*. This could be converted to the GTP-γ-S or GDP-β-S form by preincubation of the protein (15 μg ml⁻¹) with 100 μM guanine nucleotide in 2 mM EDTA for 5 min at 30 °C followed by addition of 10 mM Mg²⁺ (ref. 39). Results are expressed as mean ± s.e. of 3–6 assays.

p67^{phox}, the two other cytosolic components of the system, both contain two SH3 domains^{11,30–32} which may interact with actin-binding³³ and possibly other proteins³⁴, and both translocate to the membrane on activation³⁵. It is possible that p21^{rac1}/GDI is involved in promoting the formation of an active oxidase complex at the membrane. □

Received 5 July; accepted 2 September 1991.

- Segal, A. W. *J. clin. Invest.* **83**, 1785–1793 (1989).
- Quie, P. G., White, J. G., Holmes, B. & Good, R. A. *J. clin. Invest.* **46**, 668–679 (1967).
- Segal, A. W. *et al. New Engl. J. Med.* **308**, 245–251 (1983).
- Clark, R. A. *et al. New Engl. J. Med.* **321**, 647–652 (1989).
- Segal, A. W. & Jones, O. T. G. *Nature* **276**, 515–517 (1978).
- Cross, A. R., Jones, O. T. G., Harper, A. M. & Segal, A. W. *Biochem. J.* **194**, 599–606 (1981).
- Knoller, S., Shpungin, S. & Pick, E. *J. biol. Chem.* **266**, 2795–2804 (1991).
- Segal, A. W., Heyworth, P. G., Cookcroft, S. & Barrowman, M. M. *Nature* **316**, 547–549 (1985).
- Volpp, B. D., Nauseef, W. M. & Clark, R. A. *Science* **242**, 1295–1297 (1988).
- Nunoi, H., Rotrosen, D., Gallin, J. I. & Malech, H. L. *Science* **242**, 1298–1301 (1988).
- Leto, T. L. *et al. Science* **248**, 727–730 (1990).
- Pick, E., Kroizman, T. & Abo, A. *J. Immun.* **143**, 4180–4187 (1989).
- Shalag, D. & Pick, E. *Biochim. biophys. Acta* **1037**, 405–412 (1990).
- Bromberg, Y. & Pick, E. *Cell. Immun.* **88**, 213–221 (1984).
- Heyneman, R. A. & Vercauteren, R. E. *J. Leuk. Biol.* **36**, 751–759 (1984).
- Bromberg, Y. & Pick, E. *J. biol. Chem.* **260**, 13539–13545 (1985).

- Abo, A. & Pick, E. *J. biol. Chem.* (in the press).
- Didsbury, J., Weber, R. F., Bokoch, G. W., Evans, T. & Snyderman, R. J. *J. biol. Chem.* **264**, 16378–16382 (1989).
- Fukumoto, Y. *et al. Oncogene* **5**, 1321–1328 (1990).
- Seifert, R., Rosenthal, W. & Schultz, G. *FEBS Lett.* **205**, 161–165 (1986).
- Gabig, T. G., English, D., Akard, L. P. & Schell, M. J. *J. biol. Chem.* **262**, 1685–1690 (1987).
- Ligeti, E., Doussiere, J. & Vignais, P. V. *Biochemistry* **27**, 193–200 (1988).
- Aharoni, I. & Pick, E. *J. Leuk. Biol.* **48**, 107–115 (1990).
- Eklund, E. A., Marshall, M., Gibbs, J. B., Crean, C. D. & Gabig, T. G. *J. biol. Chem.* **260**, 13964–13970 (1991).
- Quinn, M. T. *et al. Nature* **342**, 198–200 (1989).
- Hori, Y. *et al. Oncogene* **6**, 515–522 (1991).
- Isomura, M., Kikuchi, A., Ohga, N. & Takai, Y. *Oncogene* **6**, 119–124 (1991).
- Chardin, P. *et al. EMBO J.* **8**, 1087–1092 (1989).
- Paterson, H. F. *et al. J. Cell Biol.* **111**, 1001–1007 (1990).
- Rodaway, A. R. F., Teahan, C. G., Casimir, C. M., Segal, A. W. & Bentley, D. L. *Molec. cell. Biol.* **10**, 5388–5396 (1990).
- Lomax, K. J., Leto, T. L., Nunoi, H., Gallin, J. I. & Malech, H. J. *Science* **245**, 987 (1989).
- Volpp, B. D., Nauseef, W. M., Donelson, J. E., Moser, D. R. & Clark, R. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9563 (1989).
- Drubin, D. G., Mulholland, J., Zhu, Z. & Botstein, D. *Nature* **343**, 288–290 (1990).
- Woodman, R. C. *et al. J. clin. Invest.* **87**, 1345–1351 (1991).
- Clark, R. A., Volpp, B. D., Leidal, K. G. & Nauseef, W. M. *J. clin. Invest.* **85**, 714–721 (1990).
- Pick, E. & Keisari, Y. *Cell. Immun.* **59**, 301–318 (1981).
- Diekmann, D. *et al. Nature* **351**, 400–402 (1991).
- Pick, E., Bromberg, Y., Shpungin, S. & Gadba, R. J. *J. biol. Chem.* **262**, 16476–16483 (1987).
- Hall, A. & Seif, A. J. *J. biol. Chem.* **261**, 10963–10965 (1986).

ACKNOWLEDGEMENTS. We thank the Wellcome Trust, Cancer Research Campaign, Medical Research Council, US-Israel Binational Science Foundation, the David and Natalie Roberts Chair in Immunopharmacology, Mrs J. Saltzman, Mr and Mrs J. Garbell, and Mr and Mrs L. Randall for support, D. Diekmann and S. Brill for recombinant proteins, and G. Scrase for database searches.

Phosphorylation of *c-jun* mediated by MAP kinases

Bernd J. Pulverer*, John M. Kyriakis†, Joseph Avruch†, Eleni Nikolakaki* & James R. Woodgett*‡

* Ludwig Institute for Cancer Research, 91 Riding House Street, London W1P 8BT, UK

† Diabetes Unit and Medical Services, Massachusetts General Hospital and the Department of Medicine, Harvard Medical School, Boston, Massachusetts 02114, USA

THE proto-oncogene *c-jun* is a component of the AP-1 transcription factor family involved in the mediation of nuclear events elicited by extracellular stimuli^{1–3}. The *c-jun* protein is negatively regulated by phosphorylation of residues near the carboxy terminus which are dephosphorylated in response to phorbol esters⁴. Here we identify two serine residues in the amino terminal A1 transactivation domain which are phosphorylated in response to a variety of mitogens, phorbol esters and activated *ras* (ref. 5). We present evidence that mitogen-activated protein-serine (MAP) kinases (pp54 and pp42/44) specifically phosphorylate these sites and that their phosphorylation positively regulates the transacting activity of *c-jun*. The MAP kinase enzymes pp54 and pp42/44 are regulated by tyrosine as well as serine/threonine phosphorylation^{6,7}. MAP kinase activation of *c-jun* may underlie the common stimulation of this transcription factor by mitogens, growth factors and oncogenes.

In cultured cells, *c-jun* is phosphorylated on serine and threonine residues located in tryptic peptides termed a, b, c, X and Y (refs 4, 5). a, b, and c represent mono-, di- and triphosphorylated forms of a single peptide located near the C-terminal DNA-binding domain, phosphorylation of which inhibits DNA binding⁴. These sites are specifically dephosphorylated in response to phorbol esters accounting, in part, for activation of the transcription factor⁴. Serine phosphorylation of the X and Y peptides, which have been localized to the N-terminal 87 amino acids of *c-jun* correlates with activation of the trans-activation potential of *c-jun* in response to *ras* (ref. 5). Using a combination of phosphopeptide mapping, synthetic peptides and site-directed mutagenesis we have identified serine 72 (human sequence) as the phosphorylated residue in peptide X

‡ To whom correspondence should be addressed.

and serine 62 as the residue in peptide Y. Thus, phosphorylated synthetic peptides corresponding to the predicted tryptic peptides containing serines 62 and 72 comigrate on two-dimensional electrophoresis/chromatography with peptides Y and X, respectively, from tryptic digests of *c-jun* immunoprecipitated from ^{32}P -phosphate-labelled U937 cells (Fig. 1). As the tryptic peptide corresponding to X contains a single serine residue, the phosphorylation site in this peptide is residue 72. Peptide Y contains two serine residues corresponding to residues 57 and 62 located two and seven amino acids from its N terminus. Three cycles of manual Edman degradation of phosphopeptide Y eluted from digests of *in vivo* labelled *c-jun* failed to release radioactive phosphate indicating residue 62 as the target (not shown).

Phorbol ester treatment of U937 cells induces specific phosphorylation of the X and Y peptides (and dephosphorylation of the C-terminal sites) (Fig. 1). Phorbol esters not only stimulated ^{32}P -phosphate incorporation into *c-jun* in U937 cells but also caused electrophoretic retardation of the protein (Fig. 2a). A similar effect was observed on phorbol ester treatment of HepG2 cells (Fig. 2a). The *c-jun* 'ladder' comprises 4–5 bands reminiscent of the products of *c-jun* RNA translated in reticulocyte lysates (Fig. 2b). Incubation with phosphatase collapses the ladder indicating it to be caused by phosphorylation. Reticulocyte translation of *c-jun* RNAs in which serines 62 and/or 72 were mutated to leucine abolished retardation demonstrating the effect to be due to phosphorylation of these sites (Fig. 2c, odd lanes). These data are consistent with 12-*O*-tetradecanoylphorbol-13-acetate (TPA) induction of X and Y phosphorylation and the appearance of the retarded forms of *c-jun* in immunoprecipitates (Fig. 2a). Translation of the wild-type (or mutant) *c-jun* RNA in wheatgerm extracts resulted in no electrophoretic retardation suggesting the plant extracts are

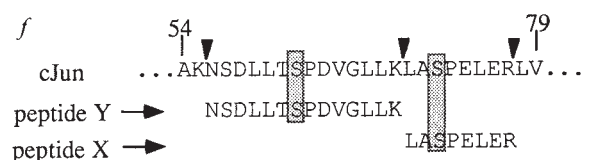
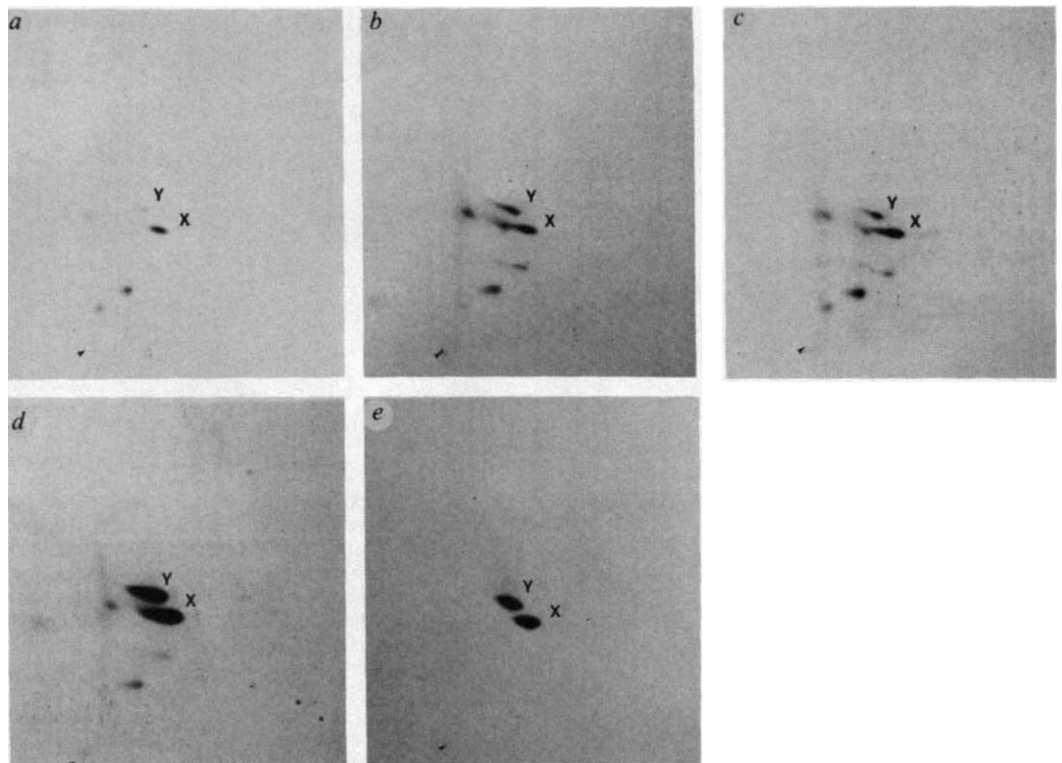
deficient in a *c-jun* kinase activity present in reticulocyte lysates (Fig. 2c, even lanes). Indeed, addition of reticulocyte lysate to the wheatgerm translated protein (together with unlabelled methionine to prevent further incorporation of radioactivity) restores the ladder effect (not shown).

The protein kinase(s) that targets these two serine residues is unknown, but in all *c-jun* proteins sequenced, residues corresponding to serine 62 and 72 are followed by prolines, prompting us to investigate whether mitogen-activated 'proline-directed' kinases are candidate *c-jun* kinases. These enzymes include pp42 and the related pp44 MAP kinases^{6,8} (also termed ERK-2 and -1, respectively⁹) as well as the recently identified pp54 MAP kinase⁷. These kinases all target serine/threonine residues preceded by proline but have discernable substrate preferences *in vitro*¹⁰ (J.A., unpublished results). In addition, the MAP kinases are regulated by tyrosine and serine/threonine phosphorylation and have been implicated in the cellular response to many growth factors^{8,10–13} and activated *ras* (S. J. Leivers and C. J. Marshall, personal communication).

To investigate the possible role of the MAP kinases in *c-jun* regulation we took advantage of the ability of wheatgerm translated *c-jun* protein to report specifically phosphorylation of the two N-terminal sites by the appearance of electrophoretically retarded forms. This assay does not detect protein kinases, such as GSK-3, that phosphorylate *c-jun* at other sites as phosphorylation at these residues has no effect on gel mobility (J.R.W., unpublished results). Using this assay it is clear that, at identical MAP-2 phosphorylating activities, pp54 and pp42/44 MAP kinases can specifically promote the mobility shift of *c-jun* with the latter kinases appearing some fivefold less potent (Fig. 3a). It should be noted that the concentration of *c-jun* protein in these extracts is very low representing <0.05% of the total

FIG. 1 Identification of *c-jun* phosphopeptides X and Y. Tryptic phosphopeptide maps of: a, *c-jun* immunopurified from resting U937 cells; b, *c-jun* immunopurified from U937 cells treated for 15 min with 160 nM TPA; c, mix of a and b; d, mix of b and e; e, synthetic peptides (*c-jun* 56–68, 69–77) phosphorylated by pp54 MAP kinase. f, Tryptic peptides encompassing serines 62 and 72. Arrowheads denote sites of tryptic cleavage.

METHODS. Human U937 cells were labelled overnight with ^{32}P -inorganic phosphate at 1 mCi ml^{-1} , treated for 15 min with or without 160 nM TPA and immunoprecipitated with a *c-jun*-specific antibody raised against the C-terminal 15 amino acids of human *c-jun*. After trypsinolysis, phosphopeptides were electrophoresed at pH 1.9 and subjected to ascending chromatography as in ref. 4. Synthetic peptides corresponding to tryptic peptide Y (NSDLLTSPDVGLLK; single-letter amino-acid code) and tryptic peptide X (LASPELER) were synthesized using Fmoc chemistry and HPLC purified. The peptides (50 pmol) were each phosphorylated by incubation with 0.5 U purified pp54 MAP kinase (see legend to Fig. 3); unincorporated ATP was removed by reverse-phase HPLC and the peptides then subjected to electrophoresis/chromatography. A total of 500 c.p.m. were spotted in each case; exposures were for 4 days at -70°C with intensifying screens.



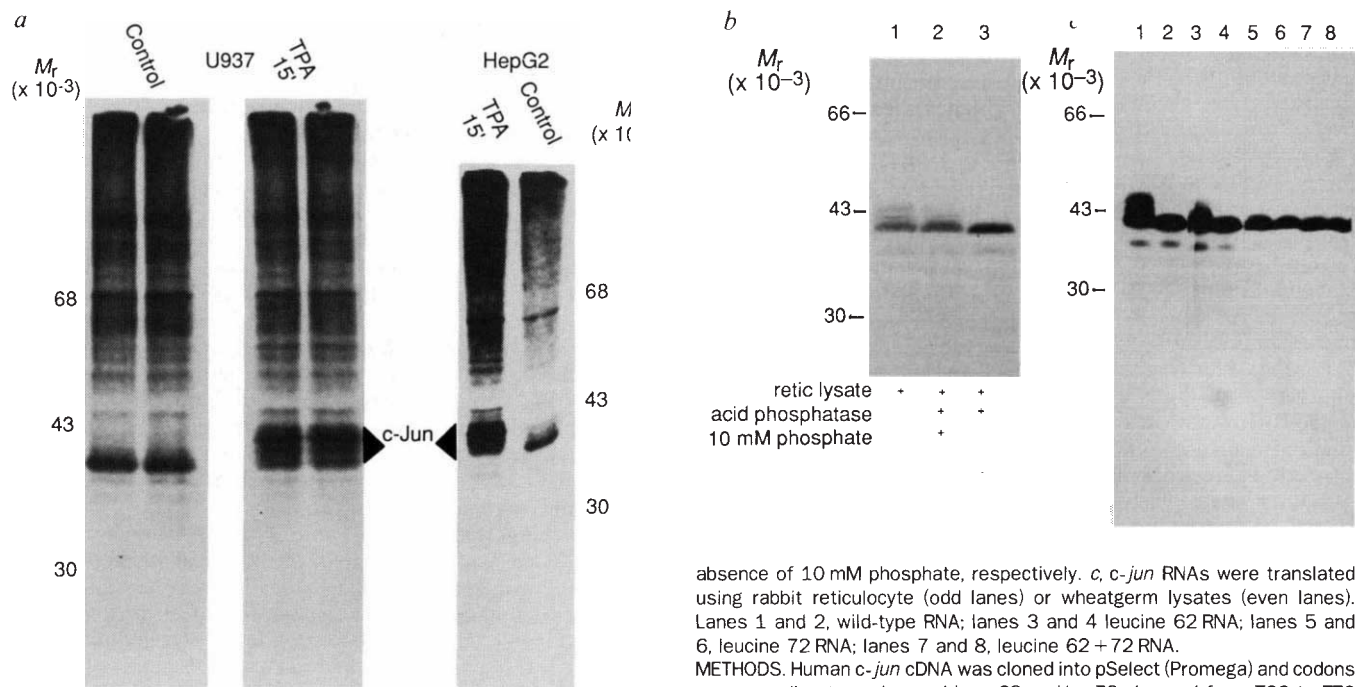


FIG. 2 Effect of N-terminal phosphorylation on gel mobility of *c-jun*. *a*, *c-jun* was immunoprecipitated from ^{32}P -phosphate-labelled U937 or HepG2 cells treated with or without 160 nM TPA as indicated. *b*, Wild-type *c-jun* RNA was translated in reticulocyte lysate and immunoprecipitated using a specific *c-jun* antiserum. In lanes 2 and 3, the immunoprecipitate was treated for 15 min with 0.5 U of potato acid phosphatase (Sigma) in the presence or

absence of 10 mM phosphate, respectively. *c*, *c-jun* RNAs were translated using rabbit reticulocyte (odd lanes) or wheat germ lysates (even lanes). Lanes 1 and 2, wild-type RNA; lanes 3 and 4, leucine 62 RNA; lanes 5 and 6, leucine 72 RNA; lanes 7 and 8, leucine 62 + 72 RNA. METHODS. Human *c-jun* cDNA was cloned into pSelect (Promega) and codons corresponding to serine residues 62 and/or 72 changed from TCG to TTG (leucine) by oligonucleotide-directed mutagenesis. The mutants and wild-type cDNA were inserted into pSP64polyA (Promega) for synthesis of RNA *in vitro*. SP6 polymerase-transcribed RNA was purified and 2 μg translated using either rabbit reticulocyte lysate or wheat germ lysate in the presence of 20 μCi ^{35}S -methionine (Amersham) according to the manufacturer's protocols (Promega). *b*, Proteins were immunoprecipitated with *c-jun* antiserum before resolution by SDS-PAGE and fluorography.

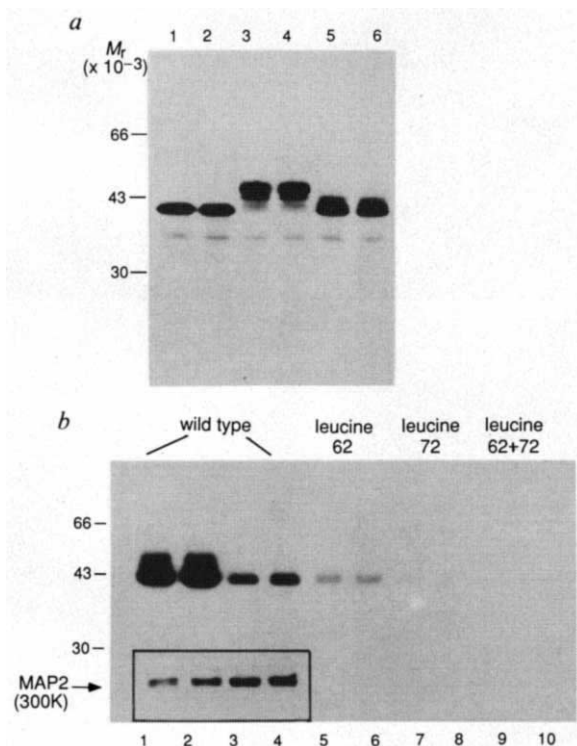
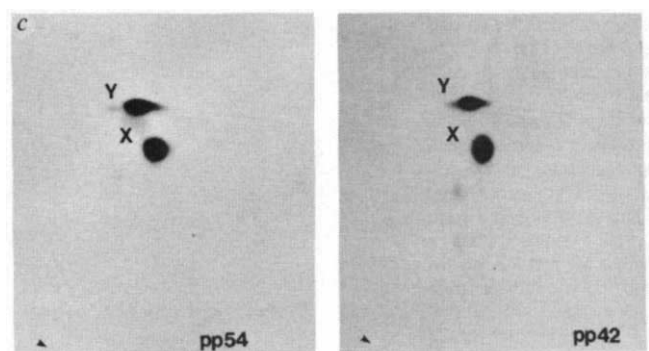


FIG. 3 MAP kinases specifically phosphorylate the two sites in the *trans*-activation domain of *c-jun*. *a*, ^{35}S -labelled *c-jun* protein translated in wheat-germ lysate was incubated with buffer (lanes 1 and 2), pp54 MAP kinase (lanes 3 and 4) or pp42/44 MAP kinases (lanes 5 and 6). *b*, Phosphorylation of bacterially synthesised *c-jun* proteins by pp54 (lanes 1 and 2) and



pp42/44 MAP kinases (lanes 3–10). Lanes 1–4 contain wild-type *c-jun*. Lanes 5 + 6, 7 + 8 and 9 + 10 contain purified *c-jun* with leucine substituted at residue 62, 72 and 62 + 72, respectively. Inset shows incorporation of ^{32}P -phosphate into microtubule associated protein-2 (MAP2) using equivalent amounts of each kinase used in lanes 1–4. *c*, tryptic phosphopeptide maps of bacterially expressed *c-jun* phosphorylated by pp54 (left) and pp42/44 (right) MAP kinases. METHODS. pp54 MAP kinase was purified from rat liver as described⁷. The pp42/44 MAP kinase preparation was isolated from 500 10-cm plates of insulin-stimulated H4 cells as described¹². This preparation contains equal amounts of ERK-1 and -2 as judged by immunoblotting⁹. For phosphorylation studies, 0.23 U pp54 and 0.33 U pp42 MAP kinase were used⁷. Wheat germ translated material (Fig. 2 legend) was phosphorylated in the presence of 1 mM ATP, 10 mM MgCl_2 , 1 mM NaVO_4 , 500 nM okadaic acid for 15 min at 30 °C. Human *c-jun* protein was expressed in pLysS bacteria using the T7 promoter containing vector pRK 171 (ref. 18) after introduction of an *Nde*I site at the initiating ATG through polymerase chain reaction-directed mutagenesis. Protein was renatured from inclusion bodies and purified by Mono Q anion exchange chromatography. Phosphorylation of bacterially expressed *c-jun* by MAP kinases for tryptic mapping was performed as above except that 100 μM [γ - ^{32}P]ATP (31,000 c.p.m. pmol^{-1}) was used.

protein. Nevertheless, the *c-jun* is quantitatively modified indicating a very high affinity of the kinase for this substrate. The low amount of *jun* protein in the translation coupled with the presence of ribonucleotides precluded assay of ^{32}P -phosphate incorporation. For this analysis we expressed wild-type and mutant *c-jun* proteins in bacteria. Purified wild-type protein was effectively phosphorylated by both preparations of MAP kinase (Fig. 3b) and two-dimensional tryptic phosphopeptide mapping demonstrated each to specifically phosphorylate the X and Y peptides (Fig. 3c). But phosphorylation of the single or double leucine 62/72 mutants was significantly reduced (Fig.

3b, lanes 5–8). Highly purified preparations of protein kinase C, protein kinase A, GSK-3, *cdc2*/cyclin, casein kinase-II, and calmodulin-dependent protein kinase-II were all incapable of phosphorylation of the X and Y peptides (not shown).

To determine the identity of the phorbol ester-stimulated *c-jun* X/Y-site kinase in U937 cells we fractionated extracts using Mono Q beads and assayed for kinase activity using purified bacterially expressed *c-jun* (Fig. 4a). TPA caused a >40-fold stimulation of a *c-jun* kinase compared with fractions from untreated cells. Immunoblotting of the most active fraction with anti-phosphotyrosine antibodies revealed a protein of relative

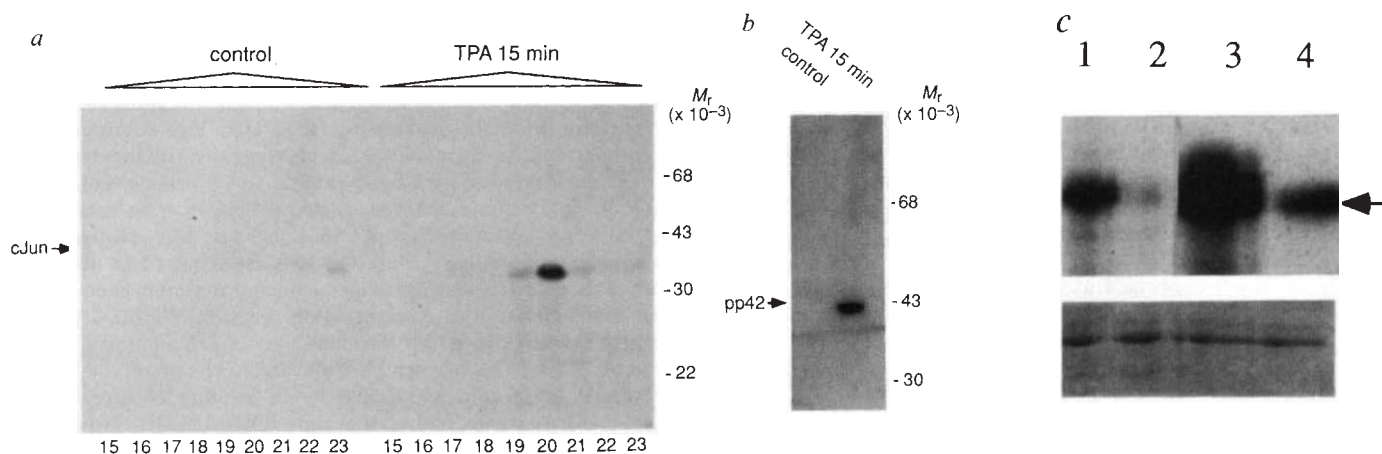


FIG. 4 Phorbol esters activate pp42 MAP kinase in U937 cells. *a*, Lysates from control and 160 nM TPA-treated U937 cells were fractionated and assayed for *c-jun* kinase activity. *b*, Anti-phosphotyrosine immunoblot of fraction 20. *c*, effect of tyrosine phosphatase on *c-jun* kinase activity. The peak fraction from the phorbol ester-treated U937 cells (lanes 1 and 2) or purified pp54 MAP kinase (lanes 3 and 4) were incubated with purified phosphotyrosyl phosphatase in the presence (lanes 1 and 3) or absence (lanes 2 and 4) of 25 mM sodium glycerophosphate. The lower panel shows Coomassie blue stained *c-jun* protein after phosphatase treatment. METHODS. U937 cells (1×10^6) were serum starved for 15 h. Half were treated with 160 nM TPA for 15 min before hypotonic lysis in 1 ml 100 μM NaVO_4 . After clarification by centrifugation, the lysates were chromato-

graphed on a Mono Q FPLC column (HR5/5), eluted with a 0–500 mM NaCl gradient and fractions assayed for *c-jun* kinase activity (see legend to Fig. 3). Fraction 20 was resolved by SDS-PAGE and immunoblotted with affinity-purified anti-phosphotyrosine antibodies: detection was with ^{125}I -labelled protein A (Amersham). For dephosphorylation, 10 U purified human lung phosphotyrosyl phosphatase was incubated with the U937 *c-jun* kinase or pp54 MAP kinase for 20 min at 30 °C in the presence or absence of 25 mM glycerophosphate as an inhibitor¹⁴. After addition of glycerophosphate to the uninhibited reactions, the kinase preparations were diluted 10-fold and assayed for *c-jun* kinase activity under initial rate conditions (arrow indicates *c-jun*).

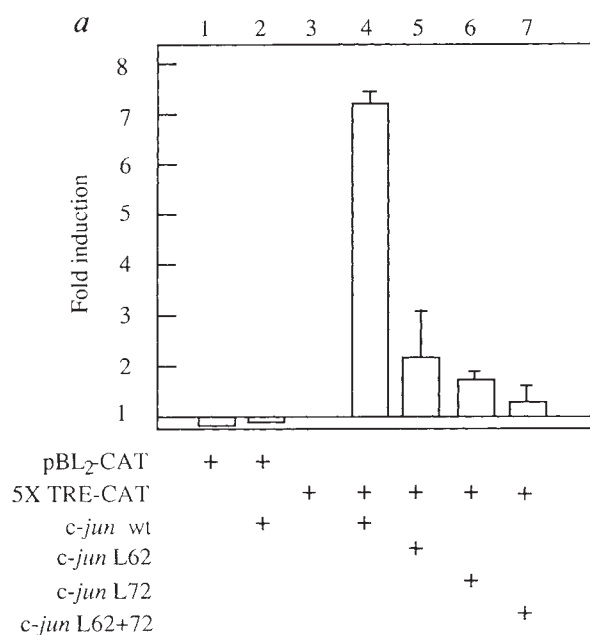
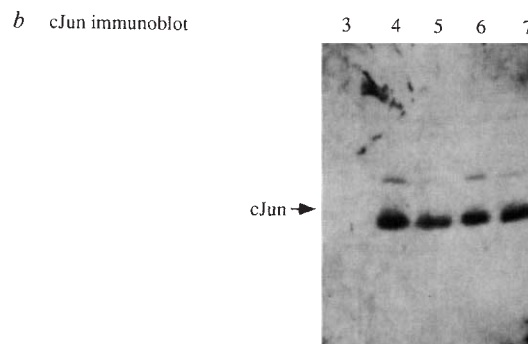


FIG. 5 Effect of N-terminal phosphorylation sites on *c-jun* activity. *a*, TRE-CAT trans-activation by various *c-jun* proteins in F9 cells as indicated.



b, *c-jun* immunoblot of transfected cell lysates. Lanes correspond to those in *a*.

METHODS. 3×10^5 (30% confluent) F9 EC cells were transfected with 3 μg CAT reporter DNA, +/- 3 μg various *c-jun* DNAs in a total of 20 μg (made up with pUC18) using CaPO_4 precipitation and cell lysates assayed for chloramphenicol acetyl transferase activity 24 h later¹⁹. Fold induction relative to TRE-CAT alone is shown. The wild-type and mutant *c-jun* cDNAs were expressed from the 3' LTR of Rous sarcoma virus. The TRE reporter plasmid is based on pBL₂CAT and contains five tandem-inverted copies of the collagenase TPA-response element upstream of the TK promoter²⁰. Transfections were in duplicate and are representative of four different experiments. The immunoblot was detected with ECL (enhanced chemiluminescence reagents) (Amersham).

molecular mass 42,000 (Fig. 4b). Further purification (5,000-fold overall) through two additional columns has established the presence of both pp42/44 MAP kinase (90% activity) and pp54 kinase (10% activity) in this TPA-induced *c-jun* kinase fraction (B.J.P. *et al.*, manuscript submitted). As MAP kinases require phosphorylation on tyrosine and threonine for activity we incubated the peak fraction of *c-jun* kinase from the U937 cells and the purified pp54 MAP kinase with a purified phosphotyrosyl protein phosphatase¹⁴ and then assayed for *c-jun* kinase activity (Fig. 4c). Preincubation with the tyrosine phosphatase inhibited subsequent serine phosphorylation of *c-jun* by 95% demonstrating that the *c-jun* kinase activity in these preparations requires tyrosine phosphorylation for activity as expected for MAP kinases.

To investigate the function of phosphorylation of serines 62 and 72 we measured *trans*-activation in F9 embryonal carcinoma cells which have undetectable levels of endogenous *c-jun*. Thus, transfection of these cells with normal *c-jun* results in *jun*-response element-dependent *trans*-activation (Fig. 5a, lanes 1–4). By contrast, transfection of plasmids containing *c-jun* in which leucine replaces serine 62 and/or serine 72, results in severe impairment of activity (Fig. 5a, lanes 5–7). The *jun* proteins were all expressed at comparable levels as judged by immunoblotting (Fig. 5b). The cells were not treated with phorbol esters thus the fraction of phosphorylated protein in these cells is low but sufficient, nonetheless, to detect changes in activity. Although we cannot completely rule out an effect of these mutations on the structural integrity of the mutant proteins, similar effects are observed with a variety of other amino-acid substitutions at the X and Y sites (data not shown). The proto-oncogene *c-jun* thus seems to be regulated both positively and negatively by protein phosphorylation events. Phosphorylation of a region proximal to the DNA-binding domain near the C terminus inhibits DNA binding⁴. By contrast, phosphorylation of two sites in the N-terminal *trans*-activation domain by a MAP kinase(s) may serve to enhance the protein's ability to stimulate gene transcription. It is noteworthy that *ras* (and phorbol esters in U937 cells) causes maximal activation of *c-jun* by promoting both dephosphorylation of the inhibitory C-terminal sites and phosphorylation of the N-terminal sites, resulting in a rapid and significant increase in the activity of this key transcription factor (Fig. 1)⁵. We also note that serines 62 and 72 are located in the A1 activation domain of *c-jun* that has been implicated in the interaction of *c-jun* with a cell type-specific inhibitor^{15–17}. Recently *ras* and *v-src* have been shown to counteract the effect of this inhibitor¹⁷. As both of these oncogenes also cause MAP kinase activation it will be of considerable interest to determine the relationship between N-terminal phosphorylation and the action of this inhibitory component. □

Domains specifying thrombin–receptor interaction

Thien-Khai H. Vu*, Virginia I. Wheaton*,
David T. Hung*, Israel Charo† & Shaun R. Coughlin*‡

* Cardiovascular Research Institute and Department of Medicine,
University of California, San Francisco, California 94143-0524, USA

† COR Therapeutics Inc., South San Francisco, California 94080, USA

PLATELET activation by the coagulation protease thrombin is central to arterial thrombosis, a major cause of morbidity and mortality. We recently isolated a complementary DNA encoding the platelet thrombin receptor. The extracellular amino-terminal extension of this seven transmembrane domain receptor contains the putative thrombin cleavage site LDPR/S which is critical for receptor activation. By replacing this cleavage site with the cleavage site for enterokinase, we have created a functional enterokinase receptor. This result demonstrates that all information necessary for receptor activation is provided by receptor proteolysis. Nanomolar enterokinase concentrations are required to activate this new receptor, in contrast to the picomolar thrombin concentrations that activate wild-type thrombin receptor. We identified a receptor domain critical for thrombin's remarkable potency at its receptor. This domain resembles the carboxyl tail of the leech anticoagulant hirudin and functions by binding to thrombin's anion-binding exosite. Our studies thus define a model for thrombin–receptor interaction. The utility of this model was demonstrated by the design of novel thrombin inhibitors based on receptor peptides.

Structure–activity studies with the cloned thrombin receptor suggested a working model for receptor activation. Thrombin cleaves its receptor at the LDPR/SFLL (single-letter amino-acid code) cleavage site (/representing the point of cleavage), unmasking a new N terminus beginning with the sequence SFLL... This new N terminus then functions as a tethered peptide ligand, binding to an undefined site in the body of the thrombin receptor, effecting receptor activation¹. The synthetic peptide SFLLRNPNDKYEPF which mimicks the unmasked N terminus is a full agonist for receptor activation¹. This working model of thrombin receptor activation predicts that all information necessary for receptor activation is provided by receptor proteolysis, and that receptor specificity is determined primarily by the substrate-recognition sequence. To test this prediction, we replaced the wild-type receptor's thrombin cleavage recognition sequence LDPR with the sequence DDDDK, the recognition site for the protease enterokinase². Wild-type and mutant thrombin receptors were expressed in *Xenopus* oocytes and receptor-mediated responses were assessed¹. Switching the receptor's cleavage recognition sites completely switched receptor specificity. Oocytes expressing the designed 'enterokinase receptor' responded to enterokinase but not to saturating concentrations of thrombin, whereas oocytes expressing the wild-type thrombin receptor responded to thrombin but not to enterokinase (Fig. 1a). By contrast, the concentration response curves to the thrombin receptor agonist peptide, which bypasses the proteolytic step to activate the receptor directly¹, were indistinguishable in oocytes expressing the enterokinase receptor or wild-type thrombin receptor. Maximal responses to agonist peptide were similar to those achieved with saturating concentrations of enterokinase or thrombin (not shown). These data strongly suggest that proteolytic unmasking of the thrombin receptor's SFLL... sequence provides all information necessary for full receptor activation, and that no additional hormone-like action of thrombin is required.

The concentration of thrombin required to elicit a half maximal response (EC₅₀) from oocytes expressing the wild-type

Received 26 June; accepted 18 September 1991.

- Bohmann, D. *et al.* *Science* **238**, 1386–1392 (1987).
- Angel, P. *et al.* *Nature* **332**, 166–171 (1988).
- Woodgett, J. R. *Semin. Cancer Biol.* **4**, 389–397 (1990).
- Boyle, W. J. *et al.* *Cell* **64**, 573–584 (1991).
- Binetruy, B., Smeal, T. & Karin, M. *Nature* **351**, 122–127 (1991).
- Anderson, N. G., Maller, J. L., Tonks, N. K. & Sturgill, T. W. *Nature* **343**, 651–653 (1990).
- Kyriakis, J. M. & Avruch, J. *J. Biol. Chem.* **265**, 17355–17363 (1990).
- Sturgill, T. W., Ray, L. B., Erikson, E. & Maller, J. L. *Nature* **334**, 715–718 (1988).
- Boulton, T. G. *et al.* *Cell* **65**, 663–675 (1991).
- Kyriakis, J. M. & Avruch, J. *FASEB J.* **5**, A797 (1991).
- Kyriakis, J. M., Brautigan, D. L., Ingebritsen, T. S. & Avruch, J. *J. Biol. Chem.* **266**, 10043–10046 (1991).
- Payne, D. M. *et al.* *EMBO J.* **10**, 885–892 (1991).
- Schonen-King, C., Nel, A., Williams, L. K. & Landreth, G. *Neuron* **6**, 915–922 (1991).
- Vicendo, P., Fauvel, J., Ragab-Thomas, J. M. F. & Chap, H. *Biochem. J.* **278**, 435–440 (1991).
- Bohmann, D. & Tijan, R. *Cell* **59**, 709–717 (1989).
- Baichwal, V. R. & Tijan, R. *Cell* **63**, 815–825 (1990).
- Baichwal, V. R., Park, A. & Tijan, R. *Nature* **352**, 165–168 (1991).
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. *Meth. Enzym.* **185**, 60–89 (1990).
- Darrow, A. L., Rickles, R. J. & Strickland, S. *Meth. Enzym.* **190**, 110–119 (1990).
- Chiu, R. *et al.* *Cell* **54**, 541–552 (1988).

ACKNOWLEDGEMENTS. We thank O. Nguyen for peptide synthesis, M. Fry for anti-phosphotyrosine antibodies, P. Vicendo for tyrosine phosphatase, C. Marshall for unpublished results and J. Hsuan, L. Pecorino and A. Kraft for many helpful discussions.

‡ To whom correspondence should be addressed.

TABLE 1 The hirudin-like domain of the thrombin receptor

(a) Alanine scanning and deletion mutagenesis

Mutant	Sequence	Response (%WT at 50 pM thrombin)
WT	... S F L L R N P N D K Y E P F W E D E E K N E S G L ...	100
$\Delta 51-63$	... S F L L R N P N D - - - - - - - - - - - S G L ...	0
$\Delta 51-56$	... S F L L R N P N D - - - - - E D E E K N E S G L ...	0
$\Delta 57-63$	... S F L L R N P N D K Y E P F W - - - - - S G L ...	100
HIR-C	... S F L L R N P N D G D F E E I P E E Y L Q N E S G L ...	100
K51A	... S F L L R N P N D A Y E P F W E D E E K N E S G L ...	88
Y52A	... S F L L R N P N D K A E P F W E D E E K N E S G L ...	2
E53A	... S F L L R N P N D K Y A P F W E D E E K N E S G L ...	25
P54A	... S F L L R N P N D K Y E A F W E D E E K N E S G L ...	92
F55A	... S F L L R N P N D K Y E P A W E D E E K N E S G L ...	45
W56A	... S F L L R N P N D K Y E P F A E D E E K N E S G L ...	70
E57, 59, 60Q	... S F L L R N P N D K Y E P F W Q D Q Q K N E S G L ...	100

(b) Inclusion of the domain enhancer receptor peptide cleavage by thrombin

	$K_m(\mu M)$	$K_{cat}(s^{-1})$	k_{cat}/K_m
TR28-45	900	35	0.04
TR28-60	10	130	13

a, Complementary RNAs transcribed from mutant receptors were expressed in *Xenopus* oocytes and ^{45}Ca release in response to thrombin or agonist peptide was measured (Fig. 1). Residues shown in bold represent non-native amino acids introduced by site-directed mutagenesis. The data shown represent the mean fold increase in ^{45}Ca release caused by 50 pM thrombin expressed as a per cent of the response seen in oocytes expressing the wild-type (WT) receptor; 50 pM is thrombin's EC_{50} for oocytes expressing wild-type thrombin receptor. Basal ^{45}Ca release in these oocytes was about 300 c.p.m. (0%) and stimulated release about 6,000 c.p.m. (100%) in most experiments. Results shown are means of nine determinations and are representative of three separate experiments. When oocytes expressing the various receptor mutants were stimulated with the thrombin receptor agonist peptide SFLLRNPNDKYEPF (ref. 1) at 10 μM (the EC_{50} for wild-type receptor), responses were indistinguishable from those seen with wild-type receptor (data not shown). b, Receptor peptides TR28-45 (PESKATNATLDPRSFL) and TR28-60 (PESKATNATLDPRSFLRNPNNDKYEPFWEDEE) were incubated with human alpha thrombin (0.1 Uml $^{-1}$ for TR28-60, 2 Uml $^{-1}$ for TR28-45 in Tyrodes buffer, pH 7.4, 25 $^{\circ}C$) for various times and the rate of peptide hydrolysis was quantitated by separating starting material and peptide products on reversed-phase HPLC (Vydac C-18 column; 30 min, 10–55% acetonitrile gradient in water with 0.1% trifluoroacetic acid; flow-through spectrophotometer at 210 nm). Initial rates of peptide hydrolysis were used for kinetic analysis. The results shown are representative of two separate experiments.

thrombin receptor was ~ 50 pM (Fig. 1b), whereas the EC_{50} for activation of enterokinase receptor-expressing oocytes by enterokinase was about 25 nM (Fig. 1a). This 500-fold difference in potencies suggested that the thrombin receptor might contain domains in addition to the primary cleavage recognition sequence for mediating thrombin–receptor interaction. We pos-

tulated that a thrombin receptor domain resembling the carboxyl tail of hirudin (Fig. 3a) might serve such a thrombin-binding function for the following reasons. Thrombin has an extended substrate-binding surface that recognizes residues both to the N and the C side of a substrate's cleavage site³ (Fig. 3b). Part of this extended substrate-binding surface, the anion-binding

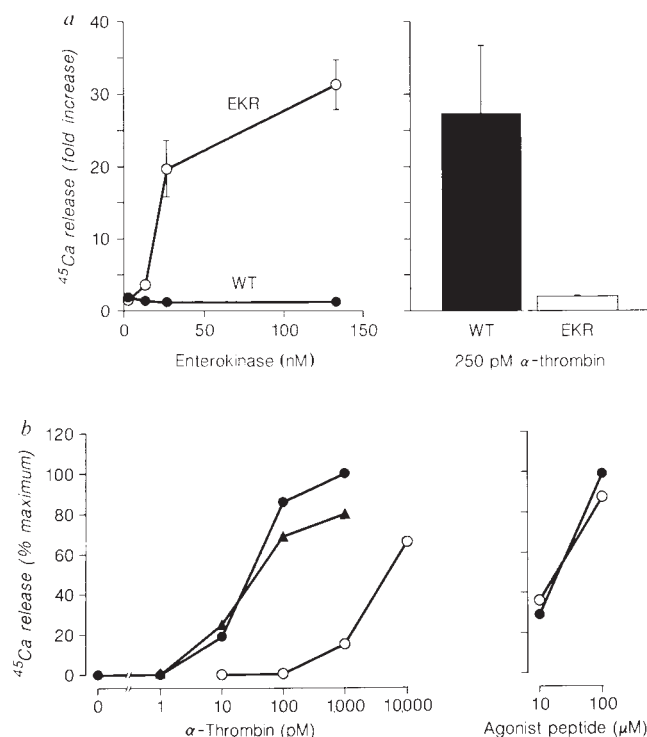


FIG. 1 Phenotypes of designed protease receptors expressed in *Xenopus* oocytes. a, Concentration response to enterokinase in oocytes expressing enterokinase receptor (EKR; open circles and column) or wild-type thrombin receptor (WT; closed symbols) is shown at left; response to a saturating concentration of thrombin is shown at right. b, Response to alpha-thrombin or to the thrombin receptor agonist peptide¹ in oocytes expressing wild-type thrombin receptor (WT; ●), thrombin receptor with its hirudin-like domain deleted ($\Delta 51-63$; ○), or the hirudin chimaera thrombin receptor (HIRC, ▲). METHODS. cDNAs encoding mutant receptors were constructed by oligonucleotide-directed mutagenesis⁹; sequences were confirmed by the dideoxy method¹⁰. The designed enterokinase receptor contained the cleavage site sequence DDDDK/S, replacing LDPR/S in the wild-type receptor. Amino-acid sequence changes in the other mutant receptors are shown in Table 1. cRNA transcribed from wild-type or mutant receptor cDNAs subcloned in pFROG was expressed in *Xenopus* oocytes and ^{45}Ca release from radiolabelled oocytes in response to alpha-thrombin, thrombin receptor agonist peptide, or enterokinase determined¹. Highly purified enterokinase free of trypsin activity was a generous gift of C. Craik, UCSF. cRNA (5ng) was injected per oocyte, pools of seven oocytes per point were used. Data shown in a are mean $\pm$ s.d. of triplicate determinations and are representative of three separate experiments. Data shown in b represent the mean of 12 replicate determinations and are representative of three separate experiments. Basal and maximal ^{45}Ca release were about 300 c.p.m. and 10,000 c.p.m. per 10 min per well, respectively, in most experiments.

exosite (Fig. 3b), is important for thrombin's ability to activate its receptor^{1,4}. The C tail of the polypeptide hirudin, a leech-derived anticoagulant which binds thrombin with remarkable avidity, binds to thrombin's anion-binding exosite⁵. Thus the presence of the hirudin-like sequence on the C side of the thrombin-cleavage site in the thrombin receptor's N-terminal extension suggested that this hirudin-like domain might mediate thrombin-receptor interaction by binding thrombin's anion-binding exosite (Fig. 3b).

FIG. 2 'Uncleavable' synthetic peptide representing the thrombin cleavage sequence and anion-binding exosite binding domain inhibits fibrinogen clotting by thrombin (open columns, thrombin plus inhibitor peptide; closed columns, thrombin plus diluent).

METHODS. The receptor peptides LDPRPFLNPNDKYEPFWEDEEKNES (TR38-64(S42P)) and LDPRPFL (TR38-45(S42P)) were synthesized by UCSF's Biomolecular Resource Center and purified by reversed phase HPLC before use. Thrombin was preincubated with peptide TR38-64(S42P) (2 μ M final concentration; open columns) or diluent (filled columns) for 5 min before addition to fibrinogen solution. Final concentrations of thrombin are shown on the abscissa. Fibrinogen clotting times were determined using a Fibro-System coagulation timer¹¹. Data shown represent mean $\pm$ s.d. of triplicate determinations (s.d. < 3% of mean). Results are representative of those obtained in three separate experiments. Unlike TR38-64(S42P), peptide TR38-45(S42P) had no inhibitory activity at concentrations as high as 300 μ M in this assay.

Deletion of the receptor's hirudin-like domain (residues 51-63) resulted in a 100-fold rightward shift of the concentration curve to thrombin, but no change in the maximal response elicited by saturating concentrations of thrombin (Fig. 1b). By contrast, the concentration response curve to agonist peptide was unaltered by the Δ 51-63 deletion (Fig. 1b and data not shown). These data demonstrate that receptor residues 51-63 are an important determinant of thrombin's potency at its receptor. Strikingly, full function could be restored to the Δ 51-63

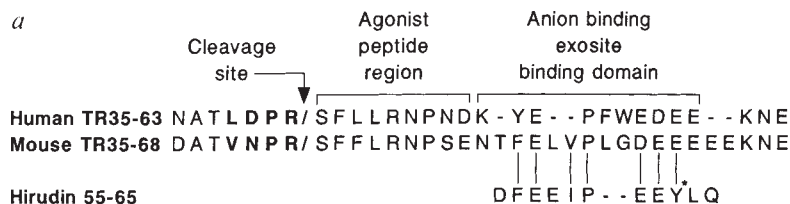
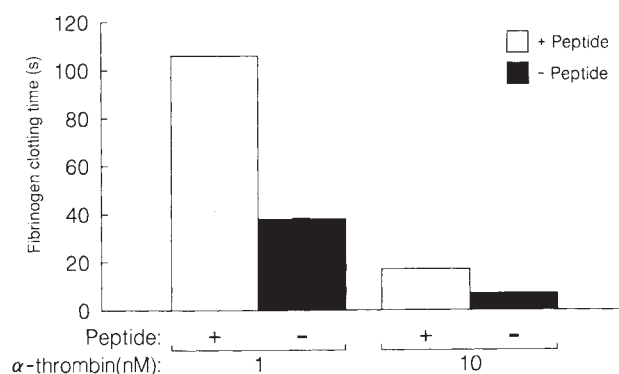
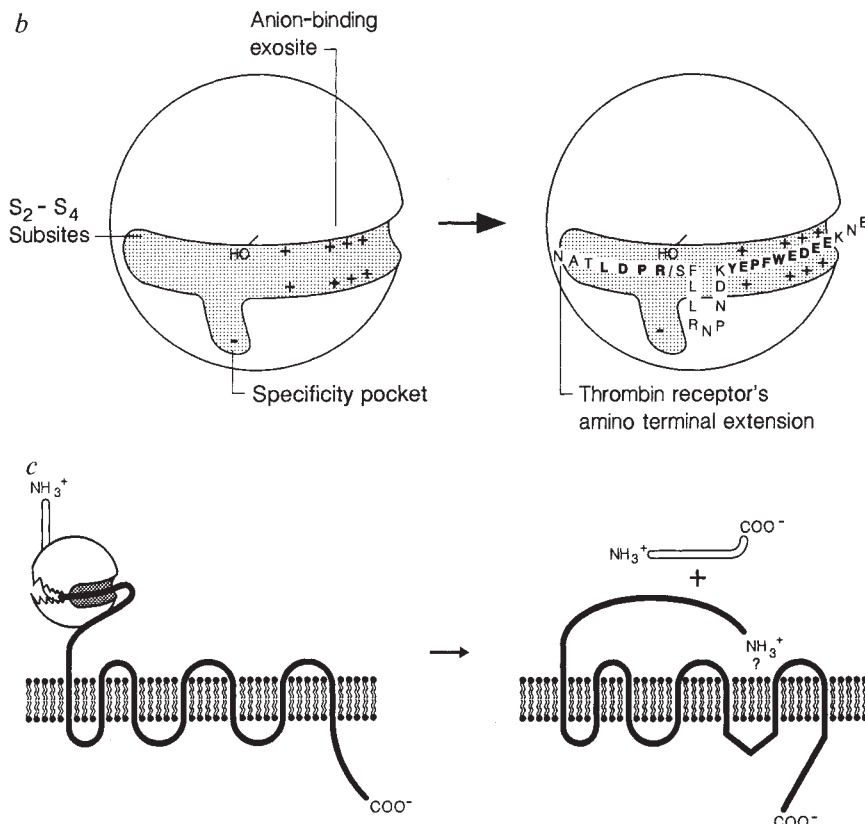


FIG. 3 Working model for thrombin-receptor interaction. **a**, Functional domains in human and murine thrombin receptor's N-terminal extension and comparison of hirudin's C tail with thrombin receptor's anion-binding exosite binding domain. Y* represents the sulphated tyrosine normally found in this position in hirudin¹². **b**, The receptor's LDPR sequence interacts with thrombin's S1-S4 subsites and active site. The receptor's YEPFWEDEE sequence binds thrombin's anion-binding exosite region. Thrombin-receptor interaction through the latter sites is a major determinant of thrombin's potency at the receptor. Data described in the text suggest that the receptor's hirudin-like anion-binding exosite binding domain and hirudin's C-tail domain share overlapping and perhaps common binding sites on thrombin. These studies in no way preclude the existence of additional receptor domains that participate in thrombin-receptor interaction. **c**, Proteolysis at the position 41/42 arginine-serine bond is sufficient for receptor activation, as shown by enterokinase's ability to activate the engineered enterokinase receptor (Fig. 1a), and by the efficacy of the agonist peptide in activating the receptor¹.



receptor by inserting the authentic hirudin C-tail domain into this position (Fig. 1b and Table a). The concentration response to thrombin of oocytes expressing this 'hirudin chimaeric receptor' (HIRC) was indistinguishable from that of oocytes expressing wild-type thrombin receptor (Fig. 1b). Thus a domain known to interact with thrombin's anion-binding exosite could substitute for the thrombin receptor residues 51–63, strongly suggesting that this receptor domain does indeed interact with thrombin's anion-binding exosite.

The effects of the $\Delta 51$ –63 deletion described above are not because of gross alterations in spatial relationships caused by the deletion (Table 1a). Rather, deletional and alanine scanning mutagenesis of the receptor's anion-binding exosite binding domain suggest that receptor residues Y52, E53 and F55 provide specific interactions important for thrombin-receptor binding, with Y52 being particularly vital (Table 1a). Thrombin-receptor residues Y52 and E53 may be analogous to hirudin C-tail residues F56 and E57, which have been shown to be critical for hirudin's binding to thrombin^{6,7} (Fig. 3). Of interest, the domain of the murine thrombin receptor (recently cloned in this laboratory) corresponding to human receptor residues 51–63 resembles hirudin's C-tail more than the human sequence (Fig. 3a). These observations suggest that the thrombin receptor and hirudin may interact with thrombin's anion-binding exosite in a similar manner, a striking example of convergent evolution.

The results described above suggested that the thrombin receptor's N-terminal extension might traverse thrombin's substrate-binding surface, the receptor's LDPR/S sequence interacting with thrombin's active site and S₁–S₄ subsites, and the YEPPWEDEE region interacting with thrombin's anion-binding exosite (Fig. 3b). This model predicts that a peptide mimicking this receptor region but rendered uncleavable by substitution of a proline for serine at the P₁' position would bind thrombin and inhibit its function. The receptor peptide LDPRPFLNPND-KYEPFWEDEEKNES ('TR38–64(S42P)') blocked all thrombin functions examined, including fibrinogen clotting (Fig. 2), and small substrate cleavage, platelet activation, and activation of oocyte-expressed native thrombin receptor (ref. 8, and data not shown). Just 2 μ M peptide was sufficient to produce dramatic prolongation of fibrinogen clotting times, even with high (1 and 10 nM) thrombin concentrations. This surprising potency of TR38–64 (S42P) contrasted with that of receptor peptide, LDPRPFL ('TR38–45(S42P)'). The latter peptide, which lacks the receptor's hirudin-like region, had little inhibitory effect in the fibrinogen clotting or other assays, even at peptide concentrations as high as 300 μ M. Indeed, kinetic analysis of the ability of these peptides to inhibit hydrolysis of small chromogenic substrates by thrombin revealed an inhibition constant (K_i) of 0.6 μ M for TR38–64(S42P), but 0.7 mM for TR38–45(S42P); occupancy of thrombin's anion binding exosite with a hirudin C-tail peptide blocked the ability of TR38–64(S42P) to inhibit thrombin⁸. Taken together, these data strongly suggest that the TR38–64(S42P) peptide occupies both thrombin's active site and anion-binding exosite, supporting the model described above (Fig. 3b).

Does the improved thrombin binding conferred by the receptor's hirudin-like anion-binding exosite binding domain translate into more efficient receptor cleavage and activation? The functional studies with receptor mutants (Fig. 1) suggest that this is so. More direct evidence was obtained using thrombin receptor-based peptides as thrombin substrates (Table 1b). Inclusion of the receptor's anion-binding exosite binding domain in a substrate peptide caused a marked increase in the efficiency of peptide cleavage by thrombin with a 90-fold improvement in the Michaelis constant (K_m) and roughly four-fold improvement in k_{cat} (Table 1b).

Taken together, these results strongly support the model of thrombin-receptor interaction illustrated in Fig. 3. Thrombin interacts with its receptor through the primary recognition sequence LDPR/S and through the anion-binding exosite binding domain sequence KYEPFWEDEE (Fig. 3a, b). Thrombin then cleaves its receptor at the LDPR/S site (Fig. 3c); the results with the enterokinase receptor (Fig. 1a) strongly suggest that this cleavage event is sufficient for receptor activation. We confirmed the utility of this model by designing a thrombin inhibitor based on a receptor peptide. This inhibitor blocked both thrombin-induced fibrinogen clotting and platelet activation; the potential of pharmaceuticals developed from this receptor-based peptide remains to be explored. □

Received 12 July; accepted 29 August 1991.

1. Vu, T.-K. H., Hung, D. T., Wheaton, V. I. & Coughlin, S. R. *Cell* **64**, 1057–1068 (1991).
2. Light, A. & Janska, H. *Trends biochem. Sci.* **14**, 110–112 (1989).
3. Bode, W. et al. *EMBO J.* **8**, 3467–3475 (1989).
4. Jakubowski, J. A. & Maragone, J. M. *Blood* **75**, 399–406 (1990).
5. Rydel, T. J. et al. *Science* **249**, 277–280 (1990).
6. Mao, S. J., Yates, M. T., Owen, T. J. & Krstenansky, J. L. *Biochemistry* **27**, 8170–8173 (1988).
7. Krstenansky, J. L., Owen, T. J., Yates, M. T. & Mao, S. J. *J. med. Chem.* **30**, 1688–1691 (1987).
8. Liu, L.-W., Vu, T.-K. H., Esmon, C. T. & Coughlin, S. R. *J. Biol. Chem.* (in the press).
9. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. *Meth. Enzym.* **154**, 367–382 (1987).
10. Sanger, F., Niklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
11. Biggs, R. *Haemostasis and Thrombosis* 2nd edn 722 (Blackwell Scientific, Oxford, 1976).
12. Niehrs, C., Huttner, W. B., Carvallo, D. & Degryse, E. *J. Biol. Chem.* **265**, 9314–9318 (1990).

ACKNOWLEDGEMENTS. We thank J. W. Fenton II (Albany Medical College, Albany, New York) for the generous gift of purified alpha-thrombin, R. Scarborough (COR Therapeutics) for peptides TR28–45 and TR28–60 and H. Bourne for comments. This work was supported by the NIH and by California's TRDRP. S.R.C. is an Established Investigator of the American Heart Association.

GUIDE TO AUTHORS

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions of the text, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in other offices, depending on the subject matter, and where necessary be sent between offices by overnight courier. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, who are chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive galley proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check their proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words, written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few subheadings of two or three words. **Letters** are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the commentary editor in the form of a one-page synopsis. Commentaries are normally between one and four pages of *Nature*.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the

News and Views editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words and 5 references.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and three copies are required, each accompanied by artwork. If photographs are included, four sets of originals are required; for line drawings, one set of originals and three good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Revised and resubmitted manuscripts should be clearly marked as such and labelled with their manuscript numbers.

Titles say what the paper is about with the minimum of technical terminology and in fewer than 80 characters in the case of Articles and Letters. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Abstract should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, an abstract should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs. Numbering of sequences should be in the left-hand margin only, with a single space between rows.

Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent the publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference, including unreferenced abstracts, should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1983–85). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used except for changed addresses.

Acknowledgements are brief, grant and contribution numbers are not allowed.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1137 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax (VAT). □

Growing human B lymphocytes in the CD40 system

Jacques Banchereau and Françoise Rousset

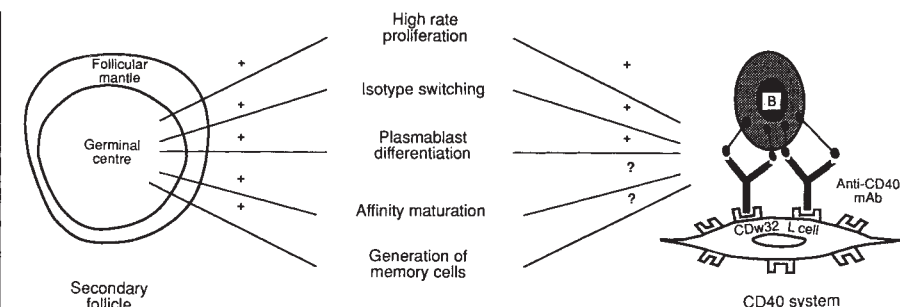
One approach to generating human monoclonal antibodies relies on the development of culture techniques allowing the long-term growth of most B lymphocytes, a property of the CD40 system.

FOLLOWING repetitive contact with a specific antigen, vertebrates produce antibodies of increasingly higher specificity and affinity, a phenomenon called affinity maturation. Specificity, affinity and safety represent highly desirable characteristics of therapeutic agents and, hence, many hopes have been given to the use of antibodies as drugs. Accordingly, immunoglobulins (Igs) isolated from human blood have proved to be extremely efficient in several clinical settings. These sources, however, have obvious limitations, such as limited availability and possible contamination by infectious agents.

The development of monoclonal antibody (mAb) technology represented a considerable achievement and resulted in numerous applications. However, in the area of immunotherapy, rodent mAbs have proved to be of limited use because of their strong immunogenicity in man. A lack of success in establishing technologies that allow the successful isolation of human hybridomas has led to the development of alternate technologies. Among these, the development of humanized rodent antibodies through the grafting of rodent mAb complementarity determining regions into an appropriate human V_H framework has already proved to be useful¹. Another approach was represented by the development of technologies allowing the long-term growth of human B lymphocytes and, most desirably, of *in vitro* techniques allowing the maturation of naive B cells into memory B cells capable of producing high-affinity, high-specificity antibodies following somatic mutations and antigen selection.

CD40 system

Monoclonal antibodies to the CD40 antigen were identified for their ability to enhance DNA replication of B cells costimulated with anti-Ig reagents². The CD40 antigen is a member of a new superfamily of membrane proteins, defined by the presence of cysteine-rich motifs, which contains the low-affinity nerve growth factor receptor, the two receptors for tumour necrosis factor (TNF) and a recently identified surface molecule that can mediate apoptosis^{3,4}. More recently, B cells were found to



Comparison of the biological properties of the CD40 system with those of a germinal centre from secondary follicles.

engage in sustained DNA replication when incubated with both the mouse fibroblastic Ltk⁻ cell line, which had been transfected with human Fc receptor (FcγRIICDw32), and mAbs to CD40⁵. This resulted in a 3–5-fold increase of input B cells within a culture period of 10–12 days. The cultures were kept for up to three weeks and provided regular enrichment of the culture medium. As resting B cells barely responded to soluble anti-CD40 mAbs, crosslinking of the CD40 antigen appeared to be necessary for B-cell proliferation. Furthermore, it is likely that the fibroblast cell line plays an important, yet undeciphered, role as immobilizing the antibody on the culture matrix gave only poor results. Under these conditions, B cells produced low levels of IgM, IgG and IgA, but no IgE.

Many different recombinant cytokines (IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, and IL-10; TNFα and TNFβ; Interferon-α (IFN) and IFNγ; and granulocyte macrophage colony-stimulating factor) and their combinations have been tested for their ability to enhance the observed proliferation. Interleukin-4 induced a strong stimulation of B-cell proliferation and allowed long-term B-cell growth. The B-cell lines grew in tight clumps. Within five weeks, the total B-cell population had expanded 100–1,000 fold. Such B-cell lines could be kept for up to 10 weeks and could be frozen and thawed. Cell lines could be generated from B cells isolated from tonsils, spleen, blood and cord blood. Furthermore, both naive sIgD⁺, sIgM⁺ and isotype-committed sIgD⁺ B cells could be induced to long-term growth. Long-term growing B cells failed to express Epstein-Barr virus

nuclear antigen 2 as determined by polymerase chain reaction or immunofluorescence analysis. They were strictly dependent on the presence of the anti-CD40 antibody and IL-4, as removal halted cell proliferation and subsequently resulted in cell death.

Upon infection with Epstein-Barr virus (EBV), these B-cell lines became independent of any factor and grew autonomously. The combination of CDw32 L cells, anti-CD40 and IL-4 activated most B cells and at least half of them entered into DNA synthesis. Limiting dilution studies have shown that 20–30 per cent of resting B cells can generate clones of 50–500 cells within two weeks. Interleukin-1α and IFNγ were found to enhance slightly the IL-4-induced proliferation of anti-CD40-activated B cells⁶. Both viral and human IL-10 induced a strong stimulation of B-cell proliferation and appeared to be as potent as IL-4 over a 10-day culture period, a time after which proliferation slowed down and then stopped completely after two to three weeks (F. Rousset *et al.*, submitted). B-cell cultures in IL-10 differed microscopically from those in IL-4 in that loose aggregates were observed early on; these decreased with time giving rise to cultures composed mostly of large single cells. At this stage, B cells have lost most of their B-cell-specific antigens, such as CD19 and CD20. The combination of IL-4 and IL-10 resulted in a 60–100-fold expansion of viable B cells over a two-week period.

In the presence of IL-4, B cells secreted IgM and IgG but, most notably, these cells secreted quite significant amounts of IgE as a result of induced isotype switching⁶. Interestingly, in contrast to pro-

liferation, induction of switching did not require the presence of the CDw32 L cell⁷. Combinations of IL-2 and IL-4 resulted in the production of large amounts of IgM and IgA, whereas IL-2 did not alter IL-4-induced proliferation. It is not known whether the observed IgA production resulted from isotype switching or enhanced differentiation of isotype-committed cells. The supernatant of B cells cultured in the presence of IL-10 contained very large amounts of IgM, IgG and IgA, but no IgE. Intracytoplasmic staining for immunoglobulins was positive for virtually all cells and electron microscopy demonstrated the presence of an extremely well-developed endoplasmic reticulum with cisternae, all being characteristic features of plasma cells (F. Rousset *et al.*, manuscript in preparation). Cells cultured in the presence of IL-10 and IL-4 secreted large amounts of the four isotypes. Naive IgM⁺, IgD⁺ B cells were found to produce IgA when stimulated in the CD40 system in the presence of both IL-10 and transforming growth factor- β , most likely as a consequence of isotype switching (T. Defrance *et al.*, submitted).

CD40 system: an *in vitro* germinal centre?

Germinal centres of lymphoid organ secondary follicles are thought to represent the microanatomical structure where antigen and T-cell-dependent maturation of B cells occurs⁸. Five separate events are supposed to take place there.

First, B cells proliferate very quickly as one B cell has been shown to give rise to several thousand daughter cells within three days⁸. Such a proliferation rate allows the clonal expansion of antigen-specific B cells and, therefore, provides a powerful amplification system allowing a quick neutralization of invading pathogens. Second, this intense proliferation is supposed to be accompanied by activation of the machinery inducing somatic mutations. While some mutations will generate antibody mutants with no affinity for their original antigen, other mutants will have an increased affinity and specificity for the eliciting antigen and these cells will be preferentially selected. Third, germinal centres are considered as being the site where B cells undergo isotype switching from IgM to either IgGs, IgAs or IgE. Fourth, a pool of these proliferating B cells also differentiate into plasmablasts, which will then migrate either to the bone marrow or to mucosa where they will mature into plasma cells. Lastly, another fraction of these cycling germinal centre B cells will revert to a small, resting state. These small B cells form the memory B lymphocyte pool, which will allow subsequent secondary responses. The follicular dendritic cells appear to play a key role in

these various events and in particular retain, for protracted periods of time, the eliciting antigen in the form of immune complexes.

Thus, it appears that the CD40 system shares many important features of germinal centres, including intense B-cell proliferation, isotype switching to IgE and IgA, and differentiation into plasmablast (see figure). The fibroblast line would mimic the follicular dendritic cells and the added cytokines would mimic germinal centre T cells. It remains to be determined whether and how the CD40 system can be manipulated to allow somatic mutations to occur and selection by a desired antigen.

CD40 system and human monoclonal antibodies

The combination of the CD40 system and EBV results in enhanced and sustained proliferation of resting B cells. Limiting dilution studies indicated that close to 10 per cent of circulating B cells could be transformed to generate cell lines. Transformed cell lines appear to produce IgG and IgA, whereas B cells transformed with EBV alone appear to induce the generation of B-cell lines producing mostly IgM antibody. In preliminary experiments, we have been able to isolate B-cell clones secreting IgG autoantibodies of predetermined specificity from patients displaying such circulating autoantibodies. Such lines will be the source of Ig mRNAs, which can be expressed on a large scale in appropriate heterologous hosts. Finally, the development of combinatorial expression libraries may make it possible to bypass EBV immortalization of long-term, human B-cell cultures⁹.

However, it has not as yet been demonstrated whether the screening of antibodies expressed in *Escherichia coli* will match and even surpass the screening of antibodies secreted by highly efficient B-cell cultures. □

Jacques Banchereau and Françoise Rousset are at Schering-Plough Corporation, Laboratory for Immunological Research, 27 chemin des Peupliers, BP 11, 69571 Dardilly Cedex, France. For more information, fill in reader service number 100.

1. Winter, G. & Milstein, C. *Nature* **349**, 293-299 (1991).
2. Clark, E. A. & Ledbetter, J. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4494-4498 (1986).
3. Itoh, N. *et al.* *Cell* **66**, 233-243 (1991).
4. Mallett, S. & Barclay, A. N. *Immunol. Today* **12**, 220-222 (1991).
5. Banchereau, J., de Paoli, P., Vallé, A., Garcia, E. & Rousset, F. *Science* **251**, 70-72 (1991).
6. Rousset, F., Garcia, E. & Banchereau, J. *J. exp. Med.* **173**, 705-710 (1991).
7. Jabara, H. H., Fu, S. M., Geha, R. S. & Vercelli, D. *J. exp. Med.* **172**, 1861-1864 (1990).
8. MacLennan, I. C. M., Liu, Y. J., Zhang, J. & Lane, P. J. *Curr. Topics Microbiol. Immunol.* **159**, 37-60 (1990).
9. Huse, W. D. *et al.* *Science* **246**, 1275-1281 (1989).

ADVERTISEMENTS

Professional Custom Services

DNA SYNTHESIS PROTEIN SEQUENCING PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services 1200 Wonderland Rd. S. Bldg. #7, Unit #2
London, Ontario, Canada N6L 1A8
CALL/FAX (519) 652-3758 ALT. FAX (519) 652-9622

Reader Service No.19

A new label....



....unique products

Whether it's 25µL of a monoclonal antibody,
50pmol of an oligonucleotide probe or
5mg of a neurochemical.

Now they've no place to hide!

AFFINITI Research Products Ltd.,
10 Cossall Industrial Estate, Ilkeston, DE7 5UG, UK

Tel: (+44/0)602 442232;
Fax: (+44/0)602 442313

Specialising in Neuroscience Research

Reader Service No.17

AUSTRAL Biologicals

NOVEL REAGENTS FOR RECEPTOR RESEARCH



• GROWTH FACTORS

- EGF
- IGF-I / IGF-II
- PDGF-AA / PDGF-BB
- acidic / basic FGF
- NGF mouse / bovine
- TNF
- TGF- β 1 / TGF- β 2

• GROWTH FACTOR RECEPTORS

- recombinant EGF / NGF / FGF / PDGF receptor extracellular domain

• POLYCLONAL AND MONOCLONAL ANTIBODIES

- ligands
- receptors

Tel: (415) 820-8390 Fax: (415) 820-6843
1-800-433-7105

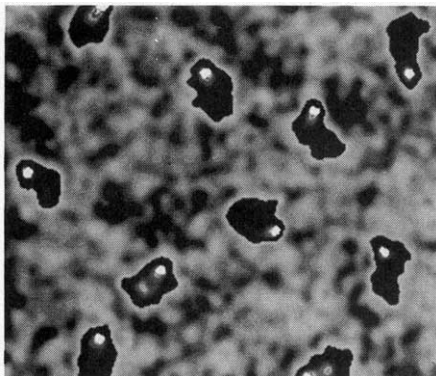
125 Ryan Industrial Ct., #207
San Ramon, California 94583

Reader Service No.20

Immunology: pick of the week

Tools for immunological research featured this week include Fab-gold immunoprobe, a wide assortment of monoclonal antibodies, a set of p53 discovery tools and a radioimmunoassay for endothelin-1.

By covalently linking 1.4-nm gold particles to Fab antibody fragments, NanoGenex has developed a new line of **Fab-gold immunoprobes** for applications in electron microscopy, light microscopy and western blot analysis (*Reader Service*



Electron micrograph of Nanogold-Fab (magnification 2,500,000).

No. 101). With a gold particle to Fab ratio that is close to 1:1, the company says its Nanogold-Fab probes offer distinct performance advantages over colloidal gold preparations, which typically use 0.2–10 variable stoichiometry and require the entire antibody fragment. The advantages include a significantly greater penetration of cells and nuclei, and an improved ability to reach sequestered antigens. The Nanogold-Fab probes can routinely detect 0.1 pg of antigen on immunoblots. However, for maximum sensitivity, the probes can be enhanced with Nanoprobes' silver developers: LI Silver for most light microscopy, electron microscopy and blotting applications and HQ Silver for electron microscopy applications where it is important to preserve internal structures. Probes are available for anti-mouse, anti-rabbit, anti-goat, anti-rat, anti-sheep and anti-biotin. Other products from NanoGenex include Nanogold-streptavidin, a full line of Nanogold-IgG probes, Nanovan (a two per cent Vanadium negative stain), non-reactive Nanogold concentrate, Monomaleimido Nanogold (for the preparation of Nanogold-Fab probes) and Monoamino Nanogold.

Immunoglobulin (Ig) isotype switching is the last stage in B-cell differentiation that involves Ig gene rearrangements. B lymphocytes within a clone can produce antibodies of different isotypes with identical antigen specificity. Janssen Biochemica has introduced the new **LIA kit for the detection of switch variants** (*Reader*

Service No. 102). The kit includes 20 strips, each of which will give a positive result after a colour development step with one of the following isotypes: $\gamma 1$, $\gamma 2a$, $\gamma 3$, λ , κ or μ . A negative control is included on each strip. Janssen says that the kit's reagents are stable for up to 12 months when stored at 4°C.

Assorted monoclonals

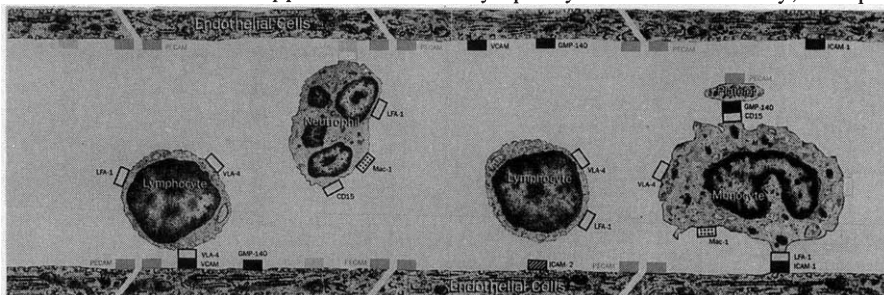
BioGenex offers a new mouse monoclonal antibody, clone G3.1, which is designed for the specific and qualitative **localization of heat-shock protein (hsp27)** on formalin-fixed, paraffin-embedded or frozen tissue sections (*Reader Service No. 103*). The high-level expression of oestrogen-regulated (27–28 kD) hsp27, especially in oestrogen receptor-positive breast tumours, has been reported to correlate with a short disease-free survival. BioGenex says that these findings suggest that hsp27 may be a useful prognostic indicator for defining a population of breast cancer patients at high risk of tumour recurrence. The antibody is sold in either a concentrated form or as a ready-to-use reagent prediluted for use with the Super Sensitive and High Performance biotin–streptavidin detection systems.

To complement the range of antibodies that are already available to the endothelial adhesion molecules, ELAM-1, ICAM-1 and PECAM, British Biotechnology has introduced a **monoclonal antibody to vascular cell adhesion molecule**, or VCAM (*Reader Service No. 104*). In response to inflammation or damage, adhesion molecules are expressed on the surface of endothelium and are thought to play a key role in the normal inflammatory process whereby leucocytes extravasate into the surrounding tissue. BBt's antibody, which is an IgG1 molecule specific for VCAM, is raised by immunizing mice with human umbilical vein endothelial cells expressing the VCAM molecule. It can be used in applications such

as immunocytochemistry, immunoprecipitation and adhesion blockade and is supplied lyophilized in 100- or 500- μ g quantities.

At the recent International Society for Analytical Cytology meeting held in Bergen, Norway, Tago Immunologicals introduced its new **F(ab')₂ TEST Reagent Family**, which includes double-staining reagents for the **detection of light chains on human B cells** by flow cytometry (*Reader Service No. 105*). Tago has combined F(ab')₂ polyclonal anti-light chain antibody with F(ab')₂ monoclonal anti-CD19 antibody to provide an accurate and precise method for the assessment of kappa and lambda light chains on the surface of B cells. As all the antibodies are F(ab')₂ fragments, Fc binding of the detecting antibodies to the cells is a problem of the past, Tago says. In addition, the F(ab')₂ TEST anti-CD19/light chain reagents have been optimized to work in whole blood on both normal and abnormal samples. In this way, researchers can use the most popular cell preparation methods and still be confident of having enough antibody even if the cells of one light chain predominate, the company says. The F(ab')₂ TEST three-vial system includes anti-CD19 R-PE blended with anti-kappa FITC, anti-CD19 R-PE blended with anti-lambda FITC, and an IgG, kappa R-PE/pre-immune FITC matched control blend. Additionally, a F(ab')₂ TEST two-vial kit consisting of anti-CD19/kappa and anti-CD19/lambda is available, and the matched control blend is also sold separately.

Cyto-Stat TiGamma A.1 is a new monoclonal antibody from Coulter designed for the **discrimination of CD3-positive lymphocytes** (*Reader Service No. 106*). The monoclonal may be of interest to researchers studying the role of T-cell receptor (TCR) gamma/delta-positive lymphocytes in autoimmunity, transplan-



BBt's monoclonal antibody to vascular cell adhesion molecule.

tation and cancer. CD3-positive lymphocytes can be discriminated on the basis of expression of either the classical CD3 cell-surface TCR alpha/beta or, the more recently identified, TCR gamma/delta. While most peripheral T lymphocytes and a large proportion of thymocytes are composed of TCR alpha/beta subunits, approximately five per cent of circulating peripheral blood mononuclear cells in normal populations are T-gamma A.1 positive. TiGamma A.1 is supplied directly conjugated to fluorescein isothiocyanate and defines a variant gene rearrangement of the gamma portion of TCR gamma/delta.

Serotec has introduced **monoclonal antibodies to the CD44, CD45 and CD62 antigens** for use in flow cytometry and fluorescence microscopy (*Reader Service No. 107*). The CD44 monoclonal recognises an antigen found predominantly on granulocytes, mature T lymphocytes and macrophages, and the antigen is a major glycoprotein on human leucocyte membranes. The CD45 monoclonal recognises the human leucocyte common antigen, which is present on all normal T and B lymphocytes, macrophages, peripheral blood lymphocytes and granulocytes, and is also present on interstitial dendritic cells. Serotec's monoclonal antibody directed against human activated platelets reacts with the 140-kD surface protein on activated platelets, megakaryocytes and activated endothelial cells. The monoclonal antibody to the CD62 antigen is designed for use in applications including flow cytometry and fluorescence microscopy.

Integri-T $\beta 1$ Integrin is the first (murine) monoclonal antibody in T Cell Diagnostics' Integri-T product line. The **monoclonal antibody binds to the human CD29 ($\beta 1$ chain) complex of the Integrin $\alpha\beta$ heterodimer** (*Reader Service No. 108*). Members of the $\beta 1$ family, also known as VLA proteins, each contain the 13-kD protein ($\beta 1$ chain), which is the common β chain of VLA (1-6) class of molecules. Integri-T $\beta 1$ Integrin may be useful in



Integri-T $\beta 1$ Integrin.

the study of infectious diseases, inflammatory diseases and certain types of cancers. It can be used alone or with ACT-T-SET VLA-1, which binds to a subset of those cells stained by Integri-T $\beta 1$ Integrin, the company says. The antibody, which is supplied in a liquid, ready-to-use form, can be used in flow cytometry and immunoprecipitation procedures and in the immunochemical staining of live, fixed and frozen cells.

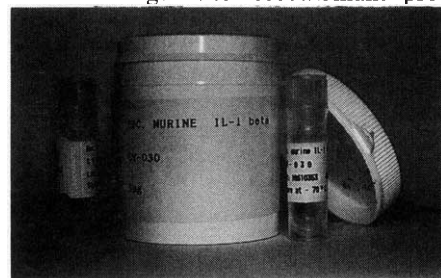
Biohit Oy is distributing the new **mouse monoclonal antibody for the neurotransmitter, calcitonin gene-related peptide (CGRP)**, which was developed by Locugenex (*Reader Service No. 109*). According to the company, the monoclonal antibody to CGRP is highly specific and eliminates the polyclonal background associated with previous polyclonal antibodies. It can be used with all commonly used cell research techniques, including fluorescence microscopy, frozen sections, vibratome sections and paraffin-embedded sections, and is available in 10- and 50- μ g quantities.

Affiniti has introduced the **PAN-Neurofilament Antibody Pack** of rabbit polyclonal antisera for the **immunochemical discrimination of neurofilament triplet proteins** (*Reader Service No. 110*). This pack contains rabbit polyclonal antisera specific for the three neurofilament proteins: NF-L (68/70 kD), NF-M (150 kD) and NF-H (200/210 kD), in addition to 50 μ g of each of the HPLC-purified proteins. The antisera can be used in immunocytochemical, western blotting

and ELISA applications, with the purified proteins serving as standards for SDS-PAGE and immunoassay quantitation. According to Affiniti, intense immunoreactivity can be obtained using routinely fixed, de-paraffinised tissue sections. Each antiserum has been extensively characterized and has been shown to detect the neurofilament proteins in rat, bovine and human tissue, Affiniti says.

Immunochemicals

A **recombinant version of murine interleukin-1 β** is now available from the Belgium-based company, Innogenetics (*Reader Service No. 111*). Murine IL-1 β , which has a molecular weight of 18 kD, has been isolated from a strain of *Escherichia coli* bearing a plasmid vector coding for the full length of the murine IL-1 β gene. Murine IL-1 β has been shown to activate thymocytes, stimulate *in vitro* bone resorption, promote growth and differentiation of B cells and stimulate the production of platelet activating factor. Innogenetics' recombinant pro-



Recombinant murine IL-1 β

duct has been purified by ammonium sulphate precipitation, anion/cation exchange chromatography and dye affinity chromatography. The company says that the product is greater than 95 per cent pure as tested by SDS-PAGE, has a low endotoxin level of less than 0.1 ng ml $^{-1}$, and has an approximate specific activity of 4×10^5 U mg $^{-1}$ as determined in an IFN-inducing bioassay with human osteosarcoma cells.

OPTIMAb from Gibco BRL is a new **monoclonal antibody production enhancer** that is designed to boost

Synthetic Oligonucleotides

Custom made probes and primers, synthesised, purified and shipped in 3 days

£3.50 per base

(All inclusive)

0.2 μ M scale (approx. 500 μ g)

Modified bases, biotin or digoxigenin labelling, data base analysis and full design facilities

Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX

Tel: XX-44-(0)71 - 326 3126

FAX XX-44-(0)71 - 733 3877

Reader Service No.56

ADVERTISEMENTS

OLIGONUCLEOTIDE SYNTHESIS

£2.95 per base

For unpurified Oligonucleotides
NO SET UP CHARGE

Digoxigenin*, Biotin, Fluorescein and Alkaline Phosphatase modification available.

Tel: (0865) 781045

Fax: (0235) 533420

British Biotechnology Limited

4-10 The Quadrant, Barton Lane,
Abingdon, Oxon OX14 3YS.

*Licensed from Boehringer Mannheim GmbH.

Reader Service No.26

Custom Peptides

\$30 a residue, 10 μ mole scale
\$60 a residue, 100 μ mole scale

Research Genetics

1-800-533-4363

In the United Kingdom 0 800 89 1393

In Canada 1-800 533 4363

FAX 205-536-9016

Reader Service No.32

monoclonal antibody yields in hybridoma culture systems by as much as 200 per cent (*Reader Service No. 112*). The enhancer, which is a protein-free, chemically-defined liquid concentrate designed as a supplement for batch hybridoma cultures, is sold as an optimized 100× concentrate of an alternate carbon source, supplemental amino acids, chemical inducer, lipids and other components. Because of its chemically-defined and protein-free formulation, OPTIMAB simplifies the purification of mAbs and minimizes serum-associated problems such as the presence of adventitious agents, Gibco says.

Immunohistochemistry

The new ImmunoDETEK Signal Generating System from Enzo uses the widely used streptavidin-biotin method for the immunohistochemical staining and detection of antibodies (*Reader Service No. 113*). Detection is based on the recognition of biotin by streptavidin, a biotin-binding protein that Enzo has linked to horseradish peroxidase. According to Enzo, streptavidin offers several advantages over avidin, namely reduced background staining due to a near neutral isoelectric point and the absence of carbohydrate side chains that are present in avidin. The kit has been designed to provide strong signals irrespective of the primary antibody used. It can be used to recognize all common classes of murine antibodies, including both IgG and IgM in paraffin-embedded tissue sections. Visualization of detection can be achieved with either aminoethylcarbazole or diaminobenzidine. Reagents are supplied concentrated in handy dropper bottles. Two sizes of kit are available, with sufficient reagents for processing either 500 or 1,000 specimens.

Vector has added four new kits to the VECTASTAIN Elite ABC family of detection systems, which are designed for use with primary antibodies made in rat, human, goat or sheep (*Reader Service No. 114*). The Elite ABC kits use the peroxidase-based detection system for immunohistochemical staining and should be of interest to pathologists, immunologists, cell biologists, anatomists and other researchers undertaking cell or tissue culture studies. Each kit consists of affinity-purified biotinylated antibody to either rat IgG (H+L), human IgG (H+L), goat IgG (H+L) or sheep IgG (H+L). In addition, each kit includes an appropriate normal serum for blocking non-specific binding sites in the tissue and the A and B reagents needed to prepare the VECTASTAIN Elite ABC pre-formed complex. Sufficient reagents are provided to prepare about 110 ml of working solution: enough to stain over 1,000 sections, smears or cytospin prepa-

rations, according to the company.

Means of measurement

Oncogene Science has developed a complete line of p53 discovery tools that are designed to aid significantly the cancer researcher in determining mutant and wild-type p53 (*Reader Service No. 115*). The line-up includes an ELISA kit for the quantitative determination of mutant p53 levels in a broad range of sample types, including cell extracts and sera. Oncogene also offers an immunohistochemistry system for staining formalin-fixed, paraffin-embedded sections for human p53 (both wild-type and mutant). Completing the package, Oncogene has added a fifth monoclonal antibody to its existing line of antibodies for the detection and characterization of both wild-type and mutant p53 of human, rat and mouse origin. This latest antibody is specific for human wild-type p53.

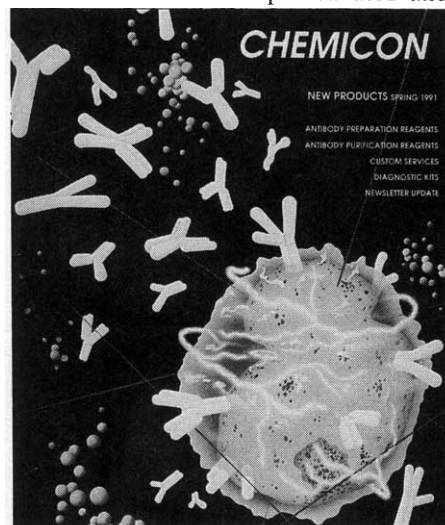
Du Pont now offers a complete radioimmunoassay kit for the measurement of endothelin-1 from plasma, serum, urine and solid tissue (*Reader Service No. 116*). Endothelin has been shown or is thought to play a role in brain, spinal cord, heart and kidney functions. It also plays a role in renal failure, myocardial infarction and diabetes mellitus. This latest RIA features a high-affinity monoclonal antibody and a simple assay procedure that requires an hour of preparation followed by an overnight incubation. The RIA has a sensitivity of 2 pg of endothelin-1 per 0.1 ml of sample. For increased sensitivity, users can follow the optional three-day assay procedure. The kit is sold in a 100-tube format with enough reagents for one standard curve and 43 duplicate samples, Du Pont says.

Catalogues/custom services

Go from synthesized peptide to purified anti-peptide antibody with the new custom anti-peptide antibody service offered by Multiple Peptide Synthesis (*Reader Service No. 117*). The service includes the following: synthesis and purification (to between 90 and 95 per cent purity) of peptide sequence, conjugation of peptide to keyhole limpet haemocyanin through a cysteine residue at the N-terminus of the peptide, immunization of two rabbits and ELISA testing of the serum, and affinity purification of the anti-peptide antibodies. Delivery times range from 13-18 weeks, MPS says. Customers receive 20 mg of purified peptide, a minimum of 30 ml crude serum, affinity purified antibodies from 5 ml serum, and a certificate of analysis indicating the reactivity of the antibody against the peptide in an ELISA. All custom research performed by MPS is confidential and the resulting peptide,

conjugate and antiserum remain the sole property of the client.

Over 100 new products are featured in the latest catalogue of immunochemicals from Chemicon (*Reader Service No. 118*). Reagents for antibody preparation include adjuvants and immunopotentiators, and a selection of carrier proteins that are also available preactivated and



More than 100 new ideas from Chemicon.

ready for conjugation. A host of conjugation reagents are listed along with their specific reaction capabilities. The catalogue also features a section devoted entirely to hybridoma production. Another new section covering reagents for antibody purification includes affinity supports — many in kit form. □

ADVERTISEMENT

DNA

Around the World

\$4.50

(no set-up)

a base

Includes:

- 500-1500 ugs
- 15-45 O.D.S.
- 24-72 hour delivery
- Amplification ready
- Gel picture
- Quality guaranteed products

We are a full service and design DNA lab.

1-800-227-0627

Fax 1-214-420-0442

BIO-SYNTHESIS